

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2003, 00:16:06 ; Search time 467 Seconds
(without alignments)
10599.341 Million cell updates/sec

Title: US-09-913-858a-1
Perfect score: 2198
Sequence: 1 actaactcaacgcgcatt.....aaaaaaaaaaaaaaaaaaaaa 2198

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: N.Geneseq_101002.*
2: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
3: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
4: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
5: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
6: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
7: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
8: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
9: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
10: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
11: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
12: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
13: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
14: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
15: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
16: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
17: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
18: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
19: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
20: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
21: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
22: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
23: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
24: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
25: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query length	DB ID	Description
1	2198	100.0	2198	21	AAA97937
2	661.6	30.1	1982	21	AAC39529
3	209.4	9.5	2069	22	AAF84907
4	209.4	9.5	2124	22	AAF84910
5	209.4	9.5	2165	22	AAF84913
6	209.4	9.5	2232	22	AAF84916
7	209.4	9.5	2808	22	AAF84906
8	209.4	9.5	2863	22	AAF84909
9	209.4	9.5	2904	22	AAF84912

10	209.4	9.5	2971	22	AAF84915	Nitrogen-inducible
11	209.4	9.5	3714	22	AAF84905	Nitrogen-inducible
12	209.4	9.5	3775	22	AAD38881	Alfalfa AP2 adapto
13	105	4.8	105	21	AAA97938	Mung bean alpha 1,
14	84.2	3.8	287	21	ABL76406	Corn tassels-derive
15	50	2.3	1461	21	AAC40119	Arabidopsis thalia
16	49.2	2.2	676	22	AA541344	CDNA encoding nove
17	49.2	2.2	1328	24	ABK11711	DNA encoding novel
18	49.2	2.2	1479	24	ABA05334	Human fucosyltrans
19	49.2	2.2	2557	24	ABA05333	Human fucosyltrans
20	47.6	2.2	358	22	AAF64428	Human fucosyltrans
21	47	2.1	34769	22	AA546775	Novel human polynu
22	46.6	2.1	34769	22	AA546774	Tumour suppressor
23	46	2.1	5931	22	AA546774	Tumour suppressor
24	45.8	2.1	6136	24	ABK40042	Tumour suppressor
25	45	2.0	612	22	AAH71471	Human chemically p
26	44.8	2.0	9814	22	ABK31429	Human cervical can
27	44.2	2.0	10020	24	ABK34293	Signal transductio
28	44.2	2.0	12405	22	AA545330	Human immune syste
29	44.2	2.0	12405	24	AA561143	Chemically pretrea
30	44.2	2.0	12405	24	ABK28169	Human gene regulat
31	43.6	2.0	626	23	ABV60941	DNA transcription
32	43.6	2.0	2036	21	AA292672	Human prostate exp
33	43.6	2.0	2170	21	AA292645	Murine alpha-1,3-f
34	43.6	2.0	113515	24	ABL34175	Murine alpha-1,3-f
35	43.4	2.0	5857	24	AA563347	Human immune syste
36	43.2	2.0	2676	21	AA292647	Chemically pretrea
37	43.2	2.0	2822	21	AA292646	Human alpha-1,3-fu
38	43	2.0	10377	22	ABJ34221	Human alpha-1,3-fu
39	42.6	1.9	9652	24	ABJ32908	Human immune syste
40	42.4	1.9	1465	21	AAC40055	Human immune syste
41	42.4	1.9	1563	21	AAC41754	Arabidopsis thalia
42	42.2	1.9	626	23	ABV60941	Arabidopsis thalia
43	41	1.9	1373	21	AACT7739	Human prostate exp
44	41	1.9	4065	21	AAFI8226	Human cancer assoc
45	41	1.9	5593	24	ABL33161	Lung cancer associ

ALIGNMENTS

RESULT 1
AAA97937 standard; DNA: 2198 BP.
ID AAA97937
AC AAA97937:
XX 26-JAN-2001 (first entry)
DT *
XX Mung bean alpha 1,3-fucosyltransferase DNA.
DE Mung bean alpha 1,3-fucosyltransferase: plant; mung bean; glycoprotein; insect;
XX alpha 1,3-fucosyltransferase: plant; mung bean; glycoprotein; insect;
KW GlcNAc-alpha1,3-fucosyl transferase; ds.
KM
KW Phaseolus aureus.
OS
XX
XX
PN WO200049153-A1.
PD
XX 24-AUG-2000.
XX
PF 17-FEB-2000; 2000WO-AT00040.
XX
PR 18-FEB-1999; 99AT-0000270.
XX
XX (ALTM/) ALTMANN F.
PA Altman F;
XX
XX
XX WPI: 2000-549274/50.
DR P-PSDB; AAB10713.
XX
XX DNA encoding plant GlcNAc-alpha-1,3-fucosyl transferase, useful for
PT recombinant production of the enzyme and recombinant glycoprotein

PR	27-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-0161359.
PR	28-JUL-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146388.	PR	28-OCT-1999;	99US-0161920.
PR	02-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161992.
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161993.
PR	04-AUG-1999;	99US-0147204.	PR	28-OCT-1999;	99US-0161993.
PR	04-AUG-1999;	99US-0147302.	PR	29-OCT-1999;	99US-0162142.
PR	05-AUG-1999;	99US-0147192.	Query Match 30.1%; Score 661.6; DB 21; Length 1982;		
PR	05-AUG-1999;	99US-0147260.	Best Local Similarity 68.1%; Pred. No. 5.5e-166;		
PR	06-AUG-1999;	99US-0147303.	Matches 987; Conservative 0; Mismatches 444; Indels 19; Gaps 4;		
PR	06-AUG-1999;	99US-0147416.			
PR	09-AUG-1999;	99US-0147493.	QY	301 AAGAGGAATGAGCAATCTATATGCTCTTGTGTGCTTGGCTGATCGCGAGATC	360
PR	09-AUG-1999;	99US-0147935.	DB	267 AAGCGAATATTGCAATTTGTATACCTCTGCGTGTGTGATGATCGTGAGATC	326
PR	10-AUG-1999;	99US-0148171.	QY	361 GCGTTTGGGTAGTGTGATATGGCCAAAACGCCGCAATGTTGACTCCCTCGCTAC	420
PR	11-AUG-1999;	99US-0148319.	DB	327 GGGTTCTGGGTGGCTCGAT-----AAAGTCGCTTGGTGTGATACCTTGACTGAT	377
PR	12-AUG-1999;	99US-0148341.	QY	421 TTCTTACCGCTCGAGCGGTGATAGGTGACGATTTGGGGTGGTTGGTGCT	480
PR	13-AUG-1999;	99US-0148565.	DB	378 TTCTTACCGCTCGAGCGGTGATAGGTGACGATTTGGGGTGGTTGGTGCT	480
PR	13-AUG-1999;	99US-0148684.	QY	481 TCTGATCGGAATCTGAATCGTATAGTTGAGAAATGTTGAGAGGAGAGATCGTTC	540
PR	16-AUG-1999;	99US-0149368.	DB	438 ATCGGATTATTACTGATAG---AGTCGCGAGAGGTGTGATGAGAGATTCAGTT	494
PR	17-AUG-1999;	99US-0149175.	QY	541 ACGTATTCGAGGGCTTTTCCAAAGAGCCTATTTTGTGTTCTGAGCTGATCAGAGTGG	600
PR	18-AUG-1999;	99US-0149426.	DB	495 ACTTACTCTAGAGATTTACTAAGATCAATTTTATCTCTTGGTGAAAGGACTTT	554
PR	20-AUG-1999;	99US-0149722.	QY	601 AAGTCGTTGGTGTGATGTAATTTGGTTAGTGGGATAGAAAGCCAGATCCGCA	660
PR	20-AUG-1999;	99US-0149722.	DB	555 CAATGCTGTTCTGTGATATTGATATTGAGATAGTTCAAGGAAACCCAGATGCTCG	614
PR	20-AUG-1999;	99US-0149722.	QY	661 TTTGGGTTACCTCAACCAAGTGAACAGCTAGCATTTCTGCATCAATGAAATCAGCAGAA	720
PR	23-AUG-1999;	99US-0149902.	DB	615 TTTGGATTAGTCAAGAAACCTGGAACCTCTTGAATATATGCTTCATTCATCAGCAG	674
PR	25-AUG-1999;	99US-0150566.	QY	721 TACTATGCTGAGAAATATTTGCCATGGCCAGAGCGGAGG---GGATATACATCGTAT	776
PR	26-AUG-1999;	99US-0150864.	DB	675 TATTATCCAAAATGATCTTGCACAGCAGCGGTGAGAGAGGTGTGATATAGTAT	734
PR	27-AUG-1999;	99US-0151065.	QY	777 GACACAGCTATCTGATGATGTCCTGTTGATATTTTCATGAGGCTGATGATAT	836
PR	27-AUG-1999;	99US-0151066.	DB	735 GACCACTAGTCTATCATATGATGTTCTGTTGATATTTTCGTGGCGGATATGATAT	794
PR	30-AUG-1999;	99US-0151303.	QY	837 GATGCAACAGTGCAGCCGAAACTGAAGCTGCTCTGACCTGCTTTCATTTCCAAATG	896
PR	31-AUG-1999;	99US-0151438.	DB	795 TATGCTCCGGTACAGCCAAAACAGAGAGACTATGTCACCTGCTTTTATTTCTAATTG	854
PR	01-SEP-1999;	99US-0151930.	QY	887 TGTGCTCGAATTTTTCCTTACAGCACTTGAGGCAATTTGAAACTCATTAATAT	914
PR	07-SEP-1999;	99US-0152363.	DB	855 TGTGCTCGAATTTTTCCTTACAGCACTTGAGGCAATTTGAAACTCATTAATAT	914
PR	10-SEP-1999;	99US-0153070.	QY	957 TGATTTTATGTTGTTGTCTACAGAGACCGTATGAGAGTGAACAAGTGAAGCCCT	1016
PR	13-SEP-1999;	99US-0153708.	DB	915 TGATTTTATGTTGTTGTCTACAGAGACCGTATGAGAGTGAACAAGTGAAGCCCT	974
PR	15-SEP-1999;	99US-0154018.	QY	1017 GAGGACTTACAAATTTAGCTTACGCTTTGAAATTTGAAATTTGAAATTTGAAATTTG	1076
PR	16-SEP-1999;	99US-0154039.	DB	975 TAAAGGATACAAATTTCAATTTGTTGGCTTTTGAAGATCTAATCAGAGAAATTTGTCACCA	1034
PR	20-SEP-1999;	99US-0154779.	QY	1077 AAAATTTTCCAAATTTTCCGTTGTTGCTGGAACCTGCTCCCTGTTGTTGTTGCTCCAAATAT	1136
PR	22-SEP-1999;	99US-0155139.	DB	1035 GAGGTTCTTTCAATCTTAGTTGCTGGGTCCGCTCCCTGTTGTTGTTGCTCCAAATAT	1094
PR	23-SEP-1999;	99US-0155486.	QY	1137 TCAGGACTTGTCTCTCTGCTGTTCAATTTTACATTTTAAAGATGAGAGATGTTGA	1196
PR	24-SEP-1999;	99US-0156599.	DB	1095 AGAAGCAATTTGCGCTGTTGGAGACTCATTTCTTACATTTAAAGACTATGAGAGATGTA	1134
PR	28-SEP-1999;	99US-0156458.	QY	1197 GTCTGTTCGAAAGCAATGAGATATCTAGCAGAAATTCGCAAGCAATATATCAATCAT	1256
PR	29-SEP-1999;	99US-0156596.			
PR	04-OCT-1999;	99US-0157117.			
PR	05-OCT-1999;	99US-0157753.			
PR	06-OCT-1999;	99US-0157865.			
PR	07-OCT-1999;	99US-0158029.			
PR	08-OCT-1999;	99US-0158232.			
PR	12-OCT-1999;	99US-0158369.			
PR	13-OCT-1999;	99US-0159293.			
PR	13-OCT-1999;	99US-0159294.			
PR	13-OCT-1999;	99US-0159295.			
PR	14-OCT-1999;	99US-0159329.			
PR	14-OCT-1999;	99US-0159330.			
PR	14-OCT-1999;	99US-0159331.			
PR	14-OCT-1999;	99US-0159637.			
PR	14-OCT-1999;	99US-0159638.			
PR	18-OCT-1999;	99US-0159584.			
PR	21-OCT-1999;	99US-0160741.			
PR	21-OCT-1999;	99US-0160767.			
PR	21-OCT-1999;	99US-0160768.			
PR	21-OCT-1999;	99US-0160770.			
PR	21-OCT-1999;	99US-0160814.			
PR	21-OCT-1999;	99US-0160815.			
PR	22-OCT-1999;	99US-0160980.			
PR	22-OCT-1999;	99US-0160981.			
PR	22-OCT-1999;	99US-0160989.			
PR	25-OCT-1999;	99US-0161404.			
PR	25-OCT-1999;	99US-0161405.			
PR	25-OCT-1999;	99US-0161406.			

```

Db 1155 GCCAGTTGCAAGAGATGATGATCTCGCAGCTAACCTGCTCTTATATAGACACT 1214
Qy 1257 GAGGTGAGATGATGAGGCTCATCTGACCTCTCAAGGCCCTTGATGATGAGCTGT 1316
Db 1215 AATATGAGAAATAGAGAGGCTCTTACAGTTCTTTCAAGGCACTGTGTATATGGCTGT 1274
Qy 1317 GCATTCATCGTCCGCTCTTTCATTCACCTTGGCCACAGTAGATGAGAGAAGAGAAA 1376
Db 1275 ACACCTCTTCTGGCGTCTGTGCATTTTCCTGGCCACAGAGGTCGAGAACAGAGAGGA 1334
Qy 1377 TAATCCAGCCTTAAAGAGAGCTCTTGCAAGT--GCATTAAGAGGCCCAAGAACCTATA 1433
Db 1335 AACCCCTAATTTCAAGAACAGACCGCTGCAATGTAGACGAGGAGATCAGACAGATTTA 1394
Qy 1434 TCATATCTATGTCAGAGAAAGAGGAGTTGATGATGATGATTCATTAAGCTAG 1493
Db 1395 TCATGTTTGTAGAGAAAGAGCGCGTTTGAATGGAATCAGCTTTTGAAGGGGTAA 1454
Qy 1494 CAATTTAACTCTGAATCTGTGAAGGCTGCTGTTTGAAGTTCAATCCCTGAATCT 1553
Db 1455 AATGTGACCTCAGAGAGCTCTAGATCTGCAAGTTCTGCGCAAGTTCAAGCTTTAAACA 1514
Qy 1554 TGTGCTGTATGAGAGACTGAAGGCTGAAGTTATAGAAGGGGAGTCTTTAAACT 1613
Db 1515 TGAGGAGAGTGTGAGAGAGAGAGGCTGGAACCTTAAAGAGACAAAGAGCTTAAAT 1574
Qy 1614 CTCAAAATATACCAATTTGCTTGACAGAGACAGACAGCTTTTATACCTTACGCTCAA 1673
Db 1575 ACATGCGATTTACCCGCTTGCCCTAACGCAACGACAGGCTTTGTACAATCTCAATTCGA 1634
Qy 1674 AGGTGATGCTGATTTCAAGAGTCACTTGAAGAACATCTTGTGCCAAGTTGAAGTCAT 1733
Db 1635 GGGAAATTCAGAGCTAGTAGTACATTTCAAAACAACCCCTTGTCTAATTTGAGGTGT 1694
Qy 1734 TTTTGTCTAG 1743
Db 1695 CTTGCTCTAG 1704

RESULT 3
AAF84907 standard; DNA: 2069 BP.
ID AAF84907
XX
AC AAF84907:
XX
DT 09-JUL-2001 (first entry)
XX
DE Nitrogen-inducible promoter for regulating foreign gene expression.
XX
KW nitrogen-inducible promoter; nitrite reductase gene; Nlr gene;
KM transgenic plant; transgene expression; ss.
XX
OS Synthetic.
XX
PN WO200125454-A2.
XX
PD 12-APR-2001.
XX
PF 02-OCT-2000; 2000WO-CO1143.
XX
PR 04-OCT-1999; 99US-0157133.
XX
PA (MEDI-) MEDICAGO INC.
XX
PI Vezina L, D'Aoust M;
XX
DR MPI; 2001-308228/32.
XX
PT Regulating foreign gene transcription in transgenic plants, comprises
PT transforming a plant (cell) with an expression construct having
PT nitrogen-inducible promoter, an open reading frame of a gene and a
PT polyadenylation site

```

```

XX
PS Claim 1; Page 30-31; 44pp; English.
XX
XX AAF84905-17 represent nitrogen-inducible promoters. They are promoters
CC of the nitrite reductase (Nlr) gene. The promoter is used for regulating
CC foreign gene transcription in transgenic organisms. The method uses an
CC expression construct having a nitrogen-inducible promoter, with(out)
CC cis-acting sequence, operably linked to the gene to be expressed and
CC modulated for transcriptional expression of the gene by addition or
CC removal of a nitrogen inducer, an open reading frame of a gene, and a
CC 3' polyadenylation signal. The method is useful for regulating the
CC transcription of transgenes in genetically modified organisms. The
CC nitrogen-inducible expression cassettes are useful for controlling
CC expression of foreign genes in plants.
XX
SQ Sequence 2069 BP; 640 A; 389 C; 335 G; 705 T; 0 other;
Query Match 9.5%; Score 209.4; DB 22; Length 2069;
Best Local Similarity 66.5%; Pred. No. 2,5e-45;
Matches 341; Conservative 0; Mismatches 151; Indels 21; Gaps 2;
Qy 1486 AGGCTAGCAATTTAACTCTGAATGCTGTAGAGCTGCTGTGTTTGAAGTTACATCC 1545
Db 170 AGATCTAGCAATTTAACTCTGAGAGCTCTTCAAGACTGCTGTCTTACGAAAGTTACGCTC 229
Qy 1546 CTGATCTTGGCTGTATGAGAACTGAAAGGCTGAAGTTATTAAGAGGGGAGTGT 1605
Db 230 CTGATATATGCTCTGTATGAGAACTGAAAGGCTGAAGTTATTAAGAGGGGAGTAA 289
Qy 1606 TTAAGCTTACAAATATATACCAATTTGCTTGACACAGACAGACAGCTTTATACCTTC 1665
Db 290 TTGAAGGTTTACAAATATATACCTTGCGGCTTGACACAGAGGCAAGCTTTATACCTTC 349
Qy 1666 AGCTTCAAGGATGCTGATTTGAGAGTCACTTGAGAGCAATCTTGCGCAAGTT 1725
Db 350 CAGTTCAAGGAGATGTTGATTTTCAAGATCCTTGAGAGCAATCTTGCGCAAGTT 409
Qy 1726 GAAGTCAATTTTGTAGATGCGCTAA-----ATGTTACTCTGCTTACCTGAAT 1778
Db 410 GAATTAATTTTGTAGATGCTTGTGAGTACCTTACATTAATGATGATCAGCAT 469
Qy 1779 AGCTTCACTTAGCTGAGCACTAGTATGAGATTTTGAAGTATGAGTATGAGTATGAG 1838
Db 470 AGCTTCTTCACTTAACTGAGATGAGATTTTGAAGTATGAGTATGAGTATGAGTATGAG 529
Qy 1839 ATGGCTT-----TATTATAGCTTCTTGTGCGCAACTGATGATGTTT 1884
Db 530 ATGGCTTGTATAGCTTACCTTACCTTGTGCGCAACTGATGAGATTTTACATTCAGAAAT 589
Qy 1885 GTATAGACATCACACTTTATTTTAACTGTTTGTAGAGTGAAGTCAATTCATTTTA 1944
Db 590 ATACATGACTTCAACATTAATTAACCCCTTTTGTAGATTAATGATGATGATTTA 649
Qy 1945 ATGCTAGTTTATGCTCTTATCTGATCATCT 1977
Db 650 ATGTTGGTGTAGTGTTTTACTGATTAAT 682

RESULT 4
AAF84910 standard; DNA: 2124 BP.
ID AAF84910
XX
AC AAF84910:
XX
DT 09-JUL-2001 (first entry)
XX
DE Nitrogen-inducible promoter for regulating foreign gene expression.
XX
KW nitrogen-inducible promoter; nitrite reductase gene; Nlr gene;
KM transgenic plant; transgene expression; ss.
XX
OS Synthetic.
XX

```


RESULT 9

AAf84912 standard; DNA; 2904 BP.

AAf84912;

09-JUL-2001 (first entry)

Nitrogen-inducible promoter for regulating foreign gene expression.

nitrogen-inducible promoter; nitrile reductase gene; NIR gene;

transgenic plant; transgene expression; ss.

Synthetic.

WO200125454-A2.

12-APR-2001.

02-OCT-2000; 2000WO-CA01143.

04-OCT-1999; 99US-0157133.

(MEDI-) MEDICAGO INC.

Vezina L, D'Aoust M;

WPI; 2001-308228/32.

Regulating foreign gene transcription in transgenic plants, comprises transforming a plant (cell) with an expression construct having a nitrogen-inducible promoter, an open reading frame of a gene and a polyadenylation site.

Claim 1; Page 34-35; 44pp; English.

AAf84905-17 represent nitrogen-inducible promoters. They are promoters of the nitrile reductase (NIR) gene. The promoter is used for regulating foreign gene transcription in transgenic organisms. The method uses an expression gene construct having a nitrogen-inducible promoter, with(out) cis-acting sequence, operably linked to the gene to be expressed and modulated for transcriptional expression of the gene by addition or removal of a nitrogen inducer, an open reading frame of a gene, and a 3' polyadenylation signal. The method is useful for regulating the transcription of transgenes in genetically modified organisms. The CC nitrogen-inducible expression cassettes are useful for the controlling expression of foreign genes in plants.

Sequence 2904 BP; 903 A; 537 C; 491 G; 973 T; 0 other;

Query Match 9.5%; Score 209.4; DB 22; Length 2904;

Best Local Similarity 66.5%; Pred. No. 2.9e-45;

Matches 341; Conservative 0; Mismatches 151; Indels 21; Gaps 2;

QY 1486 AGGCTGACAAATTAATCTGAAATGCTGTAAGGCTGCTGTTGAGTACATCC 1545
DB 909 AGATCTAGCAATTTAACTGAGAGCTCTTCAAGAGCTGCTTCTTACGAAGTTACCTCC 968
QY 1546 CTGAATCTTGCTGCTATGAGAGACTGAAAGGCTGAACTTAATAAGAGGGGAGTGTCT 1605
DB 969 CTGAATCTTGCTGCTATGAGAGAGCTGAAAGGCTGAAATTTCTAANAAGTGGGATTA 1028
QY 1606 TTTAAATCTACAAATATATCCCAATTTGCTTGACACAGACAGCAAGCTTTTATACCTTC 1665
DB 1029 TTGAAGCTTTTACAAATAATATCCCTGCGGCTTGACACAGAGCAAGCTTTTATACCTTC 1088
QY 1666 AGCTTCAAAAGTGATGCTGATTCAGAGGACACTTGAGAGCAATCTTGGCAATTTT 1725
DB 1089 CAGTTCAACGGGAGTGTGATTTTCAAGAGTCACTTGAGAGCAATCTTGGCCAAATTTT 1148
QY 1726 GAAGTCAATTTTGTGATGATGCGCTAA-----ATGATACCTCTGCTACTGATTT 1778

RESULT 10

AAf84915 standard; DNA; 2971 BP.

AAf84915;

09-JUL-2001 (first entry)

Nitrogen-inducible promoter for regulating foreign gene expression.

nitrogen-inducible promoter; nitrile reductase gene; NIR gene;

transgenic plant; transgene expression; ss.

Synthetic.

WO200125454-A2.

12-APR-2001.

02-OCT-2000; 2000WO-CA01143.

04-OCT-1999; 99US-0157133.

(MEDI-) MEDICAGO INC.

Vezina L, D'Aoust M;

WPI; 2001-308228/32.

Regulating foreign gene transcription in transgenic plants, comprises transforming a plant (cell) with an expression construct having a nitrogen-inducible promoter, an open reading frame of a gene and a polyadenylation site.

Claim 1; Page 36-37; 44pp; English.

AAf84905-17 represent nitrogen-inducible promoters. They are promoters of the nitrile reductase (NIR) gene. The promoter is used for regulating foreign gene transcription in transgenic organisms. The method uses an expression gene construct having a nitrogen-inducible promoter, with(out) cis-acting sequence, operably linked to the gene to be expressed and modulated for transcriptional expression of the gene by addition or removal of a nitrogen inducer, an open reading frame of a gene, and a 3' polyadenylation signal. The method is useful for regulating the transcription of transgenes in genetically modified organisms. The CC nitrogen-inducible expression cassettes are useful for the controlling expression of foreign genes in plants.

Sequence 2971 BP; 925 A; 549 C; 508 G; 989 T; 0 other;

Query Match 9.5%; Score 209.4; DB 22; Length 2971;

Best Local Similarity 66.5%; Pred. No. 3e-45;

Matches 341; Conservative 0; Mismatches 151; Indels 21; Gaps 2;

10-MAY-2002.

31-OCT-2001; 2001WO-CA01532.

31-OCT-2000; 2000US-244214P.

(MEDI-) MEDICAGO INC.

Veziina L, D'aoust M, Arcand F, Bliedean F;

WPI; 2002-471503/50.

Isolating and characterizing an expression regulatory sequence for
expressing recombinant polypeptides and/or RNA, comprises producing
oligonucleotide primers that amplify sequences upstream or downstream
of CDNAS

Example 2; Page 67-68; 74pp; English.

The invention relates to a method for isolating and characterising an
expression regulatory sequence for the expression of recombinant
polypeptides and/or RNA. The method comprising producing at least one
oligonucleotide primer from CDNAS of a cDNA library, where the
oligonucleotide primer allows amplification of genomic sequences upstream
or downstream of the CDNAS. The method is useful for isolating,
characterising and identifying a large number of known and unknown
promoters that are active under any desired environmental condition to
which a cell may be exposed and not just to the exemplified isolation of
promoters that are capable of expression in specific conditions. The
methods are also useful for cloning genes from any host, or from a
specific tissue with such host, from which a cDNA library may be
constructed; for the identification and isolation of analogous promoters,
signal peptides and structural genes in several species of multicellular
and unicellular organisms and as a high throughput identification system
of candidate therapeutic targets. The promoter sequences may be used to
regulate the synthesis of polypeptides. The present sequence is at least
an Ap2 adapter with nitrite reductase gene (Nir).

SQ Sequence 3775 BP; 1180 A; 658 C; 670 G; 1267 T; 0 other;

Query Match	9.58;	Score 209.4;	DB 24;	Length 3775;
-------------	-------	--------------	--------	--------------

Matches 341; Conservative 0; Mismatches 151; Indels 21; Gaps 2;

QY	1486	AGGCTTAGCAATTAAC	TACTGCTG	ATGCTG	GAAGCGCTG	CTGTTTGA	AGTTCACATCC	1545
Db	957	AGATCTAGCAATTAAC	TCTGAGAGCTCTTCA	AGACTGCTTAC	AGAAAGTTCACGTC			1016
QY	1546	CTGAATCTGGCGCTG	ATAGGAGCACTG	GAAGCGCTGA	AGCTATATAAGAGGGG	GGAGTCTCT		1605
Db	1017	CTGAATCATGTCTCG	ATAGGAAGCGCTG	AAAGCCCTCA	AAATCTAAAGGTGGCGAT	ATAA		1076
QY	1606	TTAAACCTCAAA	AAATATACCCAA	TGCTGGCTGAC	ACAGACAAGCTCTTATAC	CTTC		1665
Db	1077	TTGAAGGTTTAC	AAATATATAC	CTCGGGCTTG	TGACACAGGCAAGCTCTTATAC	CTTC		1136
QY	1666	AGCTTCAAGAGTG	ATGCTGATTT	CAGAGATCACTT	GGAAACATCTTG	TGCCAAGTTT		1725
Db	1137	CAGTTCACGGGAT	GTTGATTTTCA	CAAGTCACTT	GGAGAGCAATCTTG	TGCCAAGTTT		1136
QY	1726	GAACTTTTTTTG	TATAGCATGCGCTAA	-----ATG	TACCTGCTCA	ACCTGAAT		1778
Db	1197	GAAATATTTTTT	TGATGATATG	TGTGAGCTAC	TCTCAATTTACATGAT	CACCTAGCAT		1256
QY	1779	AGCTTCACTTAG	CTGAGCACTAG	AGTTTATG	AAATAGATATG	GCAGTGAATATG		1838
Db	1257	AGCTTTTCACTT	AACAGAGAAATGA	AGTTTATG	AAATAGATATG	GCAGTGAATATG		1316
QY	1839	ATGGCTT-----	TATTAATGCT	ATAGTTCTTG	GGCAACTCAT	TGATGTTTT		1884
Db	1317	ATGGCTTTGAT	ATAGCTCA	CCCTACTT	TGGCAACATCAT	GGGGATTTACATTC		1376
QY	1885	GTATAGACAT	CACACTTTATTTT	AAACTGTTTCT	GATAGAGTG	CAAAATCCATATTTA		1944

Db 1377 ATACATGACTTCAACCACTACTTAACCCCTTTTGTAGAATACGATGATGTCATATTTA 1436

QY 1945 ATGCTAGTTTATGCTGCTTATCTGATCATC 1977

Db 1437 ATGTTGGTGTGTAAGCTTTTACTTGATATATA 1469

```

RESULT 13
AAA97938
ID   AAA97938 standard; DNA; 105 BP.

```

PT DNA encoding plant GlcNAc- α -1,3-fucosyl transferase, useful for
PT recombinant production of the enzyme and recombinant glycoprotein
PT production -
XX
PS Claim 5; Page 36; 72pp; German.

CC This invention describes a novel DNA molecule (I), encoding a plant
CC protein with fucosyl transferase activity, GlcNAc-alpha1,3-fucosyl
CC transferase. The methods and DNA sequences are useful for production of
CC recombinant GlcNAc-alpha1,3-fucosyltransferase. The enzyme is useful for
CC the production of recombinant human glycoproteins, which are especially
CC useful in medical applications and pharmaceutical compositions. (I) can
CC be used as a probe to select GlcNAc-alpha1,3-fucosyltransferase coding
CC sequences in a sample, especially from plants or insects. This sequence
CC encodes a fragment of the mung bean (*Phaseolus aureus*)
CC alpha 1,3-fucosyltransferase protein described in the method of the
CC invention.

Sequence 105 BP; 33 A; 21 C; 20 G; 31 T; 0 other;

Query Match	4.8%	Score 105	DB 21	Length 105
Best Local Similarity	100.0%	Pred. NC	4.4e-18	
Matches 105; Conservative	0;	Mismatches	0;	Gaps 0

QY	1009	GAACCCCTGAAAGCCTTC	CAAAATTTAGCT	TCGCTTG	GAATTTG	CAATG	AGAGACAAAT	1066
Db	1	GAGCCCTGAGAGCACTCA <td>CAAAATTTAGCTT <td>AGCCTTG <td>GAATTTG <td>CAATG <td>AGAGACAAAT <td>60</td> </td></td></td></td></td>	CAAAATTTAGCTT <td>AGCCTTG <td>GAATTTG <td>CAATG <td>AGAGACAAAT <td>60</td> </td></td></td></td>	AGCCTTG <td>GAATTTG <td>CAATG <td>AGAGACAAAT <td>60</td> </td></td></td>	GAATTTG <td>CAATG <td>AGAGACAAAT <td>60</td> </td></td>	CAATG <td>AGAGACAAAT <td>60</td> </td>	AGAGACAAAT <td>60</td>	60
QY	1069	GTAACCTGAAAAATTC <td>CTTCCATCCCT <td>TGTTGCT <td>GTGGA <td>ACTGCTCCCT</td> <td>1113</td> <td></td> </td></td></td>	CTTCCATCCCT <td>TGTTGCT <td>GTGGA <td>ACTGCTCCCT</td> <td>1113</td> <td></td> </td></td>	TGTTGCT <td>GTGGA <td>ACTGCTCCCT</td> <td>1113</td> <td></td> </td>	GTGGA <td>ACTGCTCCCT</td> <td>1113</td> <td></td>	ACTGCTCCCT	1113	
Db	61	GTAACCTGAAAAATTC <td>CTTCCATCCCT <td>TGTTGCT <td>GTGGA <td>ACTGCTCCCT</td> <td>105</td> <td></td> </td></td></td>	CTTCCATCCCT <td>TGTTGCT <td>GTGGA <td>ACTGCTCCCT</td> <td>105</td> <td></td> </td></td>	TGTTGCT <td>GTGGA <td>ACTGCTCCCT</td> <td>105</td> <td></td> </td>	GTGGA <td>ACTGCTCCCT</td> <td>105</td> <td></td>	ACTGCTCCCT	105	

RESULT 14
ABL76406

ID	ABL76406 standard; cDNA; 287 bp.
XX	
AC	ABL76406;
XX	
DT	14-MAY-2002 (first entry)
XX	
DE	Corn tassell-derived polynucleotide (cdps) seq ID NO:5780.
XX	
KW	Corn; corn tassell-derived polynucleotide; cdps; hybrid breeding; CDPS; inheritance; characteristic; growth; development; disease resistance; environmental adaptability; quality; yield; molecular marker;
KW	multigene trait; plant breeding; corn tassell; gene; ss.
XX	
OS	Zea mays.
XX	
PN	US2001051335-A1.
XX	
PD	13-DEC-2001.
XX	
PF	16-APR-1999; 99US-0294093.
XX	
PR	21-APR-1998; 98US-082567P.
XX	
PA	(IALG/) IALGUDI R V.
PA	(ITOL/) ITO L Y.
PA	(SHER/) SHERMAN B K.
XX	
PI	Laijudi RV, Ito LY, Sherman BK;
XX	
DR	WPI; 2002-163647/21.
XX	
PT	Novel purified corn tassell-derived polynucleotide useful for
PT	determining altered gene expression, to recover regulatory elements and
PT	to follow inheritance of desirable characteristics through hybrid
PT	breeding programs -
XX	
PS	Claim 1; SEQ ID 5780; 201pp; English.
XX	
CC	The present sequence describes a purified corn tassell-derived
CC	polynucleotide sequence (cdps) comprising a nucleic acid sequence
CC	selected from those given in ABL7062 to ABL76833. The cdps sequences
CC	encode corn tassell-derived polypeptides (CDPS). The cdps sequences (I)
CC	can be used for determining altered gene expression, to recover
CC	regulatory elements and to follow inheritance of desirable
CC	characteristics through hybrid breeding programs. (I) are also useful
CC	in the evaluation, and alteration of desired characteristics associated
CC	with growth and development, disease resistance, environmental
CC	adaptability, quality and yield, and as molecular markers for studying
CC	inheritance of multigene traits in a plant breeding program. (I) can be
CC	used to produce a tassell-specific profile of gene transcription, a
CC	transcript image, to clone regulatory elements for use in transformation
CC	vectors, to express a polypeptide, to identify, isolate or extend
CC	identical or related corn tassell nucleic acid sequences from DNA
CC	libraries, in nucleic acid hybridisation or amplification technologies,
CC	as query sequences to determine homology of known sequences, as probe
CC	for use in Southern or Northern hybridisation, and to identify the
CC	presence of and/or to determine the degree of similarity between two
CC	(or more) nucleic acid sequences.

Query Match	3.8%	Score 84.2	DB 24	Length 287
Best Local Similarity	67.2%	Pred. No. 2.5e-12		
Matches 119: Conservative	0	Mismatches 55	Indels 0	Gaps 0
SQ Sequence 287 BP; 78 A; 57 C; 62 G; 90 T; 0 other;				

Qy	658	GCATTGGGTTACCTCAACCAAGTGAACACCTAGCACTTTCGGCATCAATGGAAATCGCA	717
Db	39	GGACTGGAGTAATCTCMTCACTTCTCTGTAGAGGTATCTCGSATTCAATGGAATCACTCT	98
Qy	718	GAACTATCTGCTGAGAACATATTTCCATGGCAAGCAGGAGGGGATTTAAACATCTGATATG	777
Db	99	CAATATATTTCAAGGAATAGTATTGATGTGCTTCGAGGAGAGAGGGTCAAGATTTGTATG	158

QY 778 ACACACAGCTATCTTCGAGATGTTCCCTGTTGGATATTTTCATGGCGTGACATGAT 834
 |||| ||| || ||||| || || ||||| || || ||||| ||||| |||||
 Db 159 ACATTCAGCCTTCTTCACACGTCGCCAGTTGGCTACTCTTCATTCATGGCGTGATATGAT 215

RESULT 15
AAC40119
ID AAC40119 standard; DNA; 1461 BP

AC AAC40119;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 27123.

KM Hybridisation assay; genetic mapping; gene expression control;

KW metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825

PR 09-MAR-1999; 99US-0123548.

PR 25-MAR-1999; 99US-0126264.

PR 01-APR-1999; 99US-0127462

PR	08-APR-1999;	99US-0128714.
PR	16-APR-1999;	99US-0128845.

PR	19-APR-1999;	99US-0130077.
PR	21-APR-1999;	99US-0130449.

PR 23-APR-1999; 99US-0130510
PR 23-APR-1999; 99US-0130891

PR 28-APR-1999; 99US-0131449;
PR 30-APR-1999; 99US-0133048

PR	30-APR-1999;	99US-0132407.
PR	04-MAY-1999;	99US-0132484

PR	05-MAY-1999;	99US-0132485,
PR	05-MAY-1999;	99US-0133485

PR	06-MAY-1999;	9905-0132487.
PR	07-MAY-1999;	9905-0132863.

PR	11-MAY-1999;	99US-0134256.
PR	14-MAY-1999.	99US-0134218

PR	14-MAY-1999;	99US-0134219
PR	14-MAY-1999;	99US-0134231

PR	14-MAY-1999;	99US-0134370.
PR	18-MAY-1999:	99US-0134768

PR	19-MAY-1999;	99US-0134941.
PR	20-MAY-1999.	99US-0135124

PR	24-MAY-1999	99US-0135639
PR	24-MAY-1999	99US-0135635

PR	25-MAY-1999;	99US-0136021.
PR	27-MAY-1999:	99US-0136393

PR	28-MAY-1999;	9905-0136782.
PR	01-JUN-1999:	9905-0137222

PR	03-JUN-1999;	99JUS-013/528.
PR	04-JUN-1999;	99JUS-0137502

PR	07-JUN-1999;	99US-0137724
PR	08-JUN-1999:	99US-0138094

PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138847

PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452

PR	16-JUN-1999;	99US-0139453.
PR	17-JUN-1999:	99US-0139492.

Mon Apr 28 09:35:59 2003

us-09-913-858a-1.rng

Page 14

Qy 1101 TGGACTGTCCCTGTGGTGTGTTGTCCTCAATATATCAGACCTTGCTCCTTCTCCT 1158
||| |||| | | | |||| | | | ||||| ||| |
Db 997 TGGTCTGTGCCAATCTATTTTGGGGCCCTTAACGTGCAGAAGACTTGTGCCCTCCGCAT 1054

Search completed: April 28, 2003, 01:22:46
Job time : 502 secs

GenCore version 5.1.4-P5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2003, 01:10:31 ; Search time 2820 Seconds
(Without alignments)
12623.291 Million cell updates/sec

Title: US-09-913-858A-1
Perfect score: 2198
Sequence: 1 actaactcaacgctgcatc.....aaaaaaaaaaaaaaaaaaaaa 2198

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estbda:*
2: em_estbhm:*
3: em_estln:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	437.6	19.9	532	13	BM143091 saj39d07
2	378.6	17.2	689	13	BM099511 BBS01_SQ
3	337.2	15.3	571	13	BJ464838
4	332.2	15.1	517	12	BG790911
5	280.2	12.7	690	14	BQ659171
6	257.8	11.7	701	14	BQ167130 WHE0946_F

7	248.8	11.3	856	10	BE412601
8	231.2 <th>10.5</th> <td>803</td> <td>12</td> <td>BF626851</td>	10.5	803	12	BF626851
9	223.8	10.2	420	9	AJ475410
10	218.8	10.0	700	9	AL503218
11	212.6	9.7	729	14	BM816799
12	208	9.5	560	14	BM737488
13	195.2	8.9	619	12	BG465531
14	191	8.7	383	13	BM443222
15	189	8.6	321	12	BG263065
16	187.2	8.5	674	13	BJ170811
17	179	8.1	844	12	BF261629
18	172.8	7.9	668	14	BQ462317
19	156.4	7.1	366	17	BH431446
20	156.4	7.1	753	17	BH418424
21	151.8	6.9 <td>735</td> <td>17</td> <td>BH423619</td>	735	17	BH423619
22	146	6.6	526	9	AL500587
23	141.8	6.5	669	17	B67847
24	136.8	6.2	772	17	BH535486
25	134.4	6.1	806	17	BH441710
26	133.8	6.1	269	10	AM695306
27	133.6	6.1	543	17	CNS00M0V
28	131.6	6.0	484	14	BM737489
29	130	5.9	496	10	AM498425
30	127.4	5.8	539	17	BH783802
31	111	5.1	529	14	BQ765453
32	110	5.0	240	9	AJ477062
33	105	4.8	604	17	AQ271924
34	105	4.8	616	17	AQ328306
35	104.6	4.8	344	17	BH440694
36	101.6	4.6	791	10	BE034958
37	101	4.6	472	12	BG241740
38	96.4	4.4	500	14	BM885010
39	89.8	4.1	568	13	BJ201098
40	86.2	3.9	235	14	BQ629003
41	84.6	3.8	433	10	AV792413
42	84.4	3.8	501	13	BI922470
43	78.6	3.6	663	17	AQ158899
44	78.6	3.6	663	17	AQ840447
45	73.8	3.4	566	17	BH776651

ALIGNMENTS

RESULT 1
LOCUS BM143091 532 bp mRNA linear EST 29-NOV-2001
DEFINITION saj39d07.y1 Gm-c1066 glycine max cDNA clone SOYBEAN CLONE ID: Gm-c1066-5150 5' similar to TR:Q9ST51 Q9ST51 FUCT C3 PROTEIN. ; mRNA sequence.

ACCESSION BM143091.1 GI:17153158
VERSION
KEYWORDS
SOURCE
ORGANISM

soybean.
Glycine max

Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eurosids I: Fabales: Fabaceae: Papilionoideae: Phaseoleae: Glycine.

REFERENCE
AUTHORS

1 (bases 1 to 532)
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corvett, V., Khana A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, T., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schenk, R., Ritter, E., Kohn, S., Shio, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800

TITLE JOURNAL
COMMENT

	Matches	433; Conservative	0; Mismatches	138; Indels	3; Gaps	1.
QY	718	GAATACTATGCTGAGAACAAATATTGCCATGGCAGACGAGGAGATATACATCGTAATG	777			
Db	1	GAATATTATTCGAGAACAAATATT---AATCGCGCTCGAGAGGGGTACCAAAATTTGTGANG	57			
QY	778	ACAACCAAGTATCATGTGGATGTGCCCTGTGGATATTTTTCATGGGCGTAGATGTATG	837			
Db	58	ACAACCAAGCTTTCTCTCGATGATGTGCCAGTTGGCTACTTTTCATGGGCTGAATATGATATC	117			
QY	838	ATGGCACCAGTGCAGCCGAAAACTGAAAGCTCTGCTTGCAGCTGCTTTCAATTTCCAAATGT	897			
Db	118	ATGCACACTGTGCTCCCAAAAGACTGAAGAGGCCCTAGCTGCAGCCTTATTTTCCAACCTGC	177			
QY	898	GGTCTCGGAATTTTCCGGTTGCAGACCTTTTGAGGCCCTTGAAAAATTCAAACATCAAAATT	957			
Db	178	GGTCCAGCAAACTTCCGTTTGCACACCCCTTGAGATGATGTGAAAGCTTTGATTTAAAAATT	237			
QY	958	GATTCCTATGTGTGTTGCACAGAGAACCGTGTGAGAACAGTAACAAAGTGAACCCCTG	1017			
Db	238	GATTCCTATGTGTGTTGCACAGAGAACCGTGTGAGAACAGTAACAAAGTGAACCCCTG	297			
QY	1018	AAGCAGTACAAATTTAGCTTAGCGTTTGAAAAATTCGAATGAGAGAGATTATGTAACCTGA	1077			
Db	298	AAGGCGTACAAATTTAGCTTAGCGTTTGAAAAATTCGAATGAGAGAGATTATGTAACCTGA	357			
QY	1078	AAATTTTCCAAATCCCTGTGTGCGAGACTGACCCCTGGTGGTGTGGCTCCCAATAT	1137			
Db	358	AAATTTTCCAAATCCCTGTGTGCGAGACTGACCCCTGGTGGTGTGGCTCCCAATAT	417			
QY	1138	CAGACACTTGGCTCCCTCCCGGTTCATTTTACATATTAAGAGATAGAGAGATGTGAG	1197			

Db	478	TCAATTGGGAAGACATTAATAACATATTTCATCATCAATCTCGATGCTTTTAAATCAATCTTG	537
QY	1258	AGGTGAGATATGAGGGTCATCTGCATCCTTCA	1291
Db	538	AGGTGAGAAGTATGATGGTCATCTGATCTTTTCA	571
RESULT 4	Bg790911	BG790911	517 bp mRNA linear EST 29-NOV-2001
LOCUS	Bg790911	sae7te02.y1 Gm-cl064	Glycine max cDNA clone GENOME SYSTEMS CLONE
DEFINITION	ID: Gm-cl064-3675 5'	similar to TR:Q9S751 Q9S751	FUCC C3 PROTEIN.
ACCESION	BG790911	:	mRNA sequence.
VERSION	BG790911.1	GI:14126473	
KEYWORDS	EST.		
SOURCE	soybean.		
ORGANISM	Glycine max		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Rosidae; Myricales; Magnoliophyta; eudicotyledons; core eudicots;		
	Fabiales; Fabaceae; Papilionoideae; Phaseoleae;		
	Glycine.		
REFERENCE	1 (bases 1 to 517)		

AUTHORS	TITLE	JOURNAL	COMMENT
Shoemaker, R., Kelm, P., Vothin, L., Erpelding, J., Corvelli, V., A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., B. Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., B. Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., R., Rilter, E., Kohn, S., Shln, T., Jackson, Y., Cardenas, M., W. R., Waterston, R., and Wilson, R.	Public Soybean EST Project	Unpublished (1999)	Contact: Shoemaker R./Public soybean EST Project Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810

Email: est@watson.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130
 South Memorial Parkway Huntsville, AL 35801 For further information
 call: (800)-533-4363 or contact via email: ccl@resgen.com
 High quality sequence stop: 421.

FEATURES

SOURCE

Location/Qualifiers
 1. 517
 /organism="Glycine max"
 /cultivar="Williams"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl064-3675"
 /clone_lib="Gm-cl064"
 /tissue_type="seedling epicotyls"
 /dev_stage="2 week old"
 /lab_host="DH10B"
 /note="Vector: pluescript II SK+ Site.1: EcoRI; Site.2:
 XhoI; The cDNA library was constructed from mRNA isolated
 from the epicotyls of 2 week old seedling for the cultivar
 Williams. The seedlings were germinated in a growth
 chamber, excised above the soil level, and the plants
 were placed in a 100 ppm solution of auxin for 24 hours
 prior to harvesting. Complementary DNA was synthesized
 from mRNA using a primer consisting of a poly(dT)
 sequence with a XhoI restriction site. EcoRI adapters
 were ligated to the blunt-ended cDNA fragments followed
 by XhoI digestion. The cDNA fragments were directionally
 cloned into the EcoRI-XhoI restriction site of the
 pluescript vector. The ligated cDNA fragments were
 transformed into DH10B host cells (GibcoBRL). This
 library was constructed in the laboratory of Dr. Randy
 Shoemaker."

BASE COUNT 147 a 88 c 110 g 172 t
 ORIGIN

Query Match 15.1%: Score 332.2; DB 12; Length 517;
 Best Local Similarity 83.7%: Pred. No. 1.1e-49;

Matches 426; Conservative 0; Mismatches 73; Indels 10; Gaps 4;

1456 GGAAGGTTGAGATGAGATCCATTACCTGAGGCTGATGCAATTTAACTGATGCTGTG 1515
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 1 GGAAGGTTGAGATGAGATCCATTACCTGAGGCTGATGCAATTTAACTGATGCTGTG 60
 1516 AAGGCTGCTGTTGTTTAACTGATCCCTGAACTTGTGCTGATVGAAGACTGAA 1575
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 61 AAGTTGCTGTTGTTTAACTGATCCCTGAACTTGTGCTGATVGAAGACTGAA 120
 1576 AGGCTGAAGTTATPAGAGGGGGGGGCTTAAACCTCAAAATATACCAATTGGC 1635
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 121 AGACCTGAATTTCTAGAGGGGGGAGATGATTTAAACCTCAAAATATACCAAGTTGGG 180
 1636 TTGACACAGACAGAGCTCTTTATACCTTACCTTCAAGGTGATGCTTACAGAGT 1695
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 181 TTGACACAGACAGAGCTCTTTATACCTTACCTTCAAGGTGATGCTTACAGAGT 240
 1696 CACTTGGAGAACATCCTTGGCCAGATTGTAAGTCATTTTGTGTACATGCGCTAAT 1755
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 241 CACTTGGAGAACATCCTTGGCCAGATTGTAAGTCATTTTGTGTACATGCGCTAAT 299
 1756 GGTACCTGCTGCTAACCGCATTTAG--CTTACCTAGCTGAGCACTACTAGAGTTTAG 1813
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 300 GGTACCTGCTGCTAACCGCATTTAG--CTTACCTAGCTGAGCACTACTAGAGTTTAG 355
 1814 GAATGAGTATGAGCATGATATGCGATGCTTTATTTATGCTAGTTCTTGGCCAACTC 1873
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 356 GGAATGAGTATGAGCATGATATGCGATGCTTTATTTATGCTAGTTCTTGGCCAACTC 415
 1874 ATTGATGTTTGTATTAACATCACAATTAAATTTAACTGTTTCTGT--AGAACTG 1930
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 416 ATTGATGTTTGTATTAACATCACAATTAAATTTAACTGTTTCTGT--AGAACTG 1930
 1931 CAATCCATATTTAATGTTAGTTTACT 1959
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 476 AGTTCCATATTTAATGCTTAGTTTACT 504

RESULT 5

BO659171/c

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BO659171 690 bp mRNA linear EST 15-JUL-2002
 HD01K22w HD Hordeum vulgare cDNA clone HD01K22 3-PRIME, mRNA
 sequence.
 BO659171
 BO659171.1 GI:21800304
 EST.
 Hordeum vulgare.
 Hordeum vulgare.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
 ; Triticeae; Hordeum.
 1 (bases 1 to 690)
 Michalek,W., Weschke,W., Pleissner,K.-P. and Graner,A.
 EST sequencing and analysis in barley
 Unpublished (2000)
 Contact: Stein Nils
 Molecular Markers Group, Department Genbank
 Institute of Plant Genetics and Crop Plant Research (IPK)
 Corrensstr. 3, 06466, Gatersleben, Germany
 Tel: 039482-5522
 Fax: 039482-5595
 Email: stein@ipk-gatersleben.de
 Insert length: 690 Std Error: 0.00
 Plate: 1 row: K column: 22
 Seq primer: T7.
 Location/Qualifiers
 1. 690
 /organism="Hordeum vulgare"
 /cultivar="Golden Promise"
 /db_xref="GABI:132513"
 /db_xref="taxon:4513"
 /clone="HD01K22"
 /clone_lib="HD"
 /tissue_type="callus"
 /dev_stage="callus (5-10 mm in diameter)"
 /lab_host="XLI0-Gold"
 /note="Vector: pluescript SK+; Site.1: EcoRI (5'-end of
 cDNA); Site.2: XhoI (3'-end of cDNA); Due to a cloning
 artefact caused by the kit, in most cases the EcoRI site
 is NOT present, as well as the EcoRI adapter used for
 cloning. To excise the insert, restriction sites upstream
 EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also
 due to the cloning system used Blue/white selection for
 recombinants is not 100% reliable. Average insert size is 1
 kb"

FEATURES

SOURCE

BASE COUNT 210 a 142 c 122 g 216 t
 ORIGIN

Query Match 12.7%: Score 280.2; DB 14; Length 690;
 Best Local Similarity 68.1%: Pred. No. 1.7e-40;

Matches 390; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

1172 ATATTAAAGATGAGAGATGCTGAGTCTGTTGGCAAGACCATGATATCTAGAGAAA 1231
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 690 ACATCAAGAGACTGATGATGATGATCTATTTCAGTTGGCAAGACATGAACATATTCATCA 631
 1232 ATCCGAAGCATATATATCAATCATTTAGAGTGAAGTATGAGGCTGCATCTGACTTCA 1291
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 630 ATCCGATGATCTTTAATCAATCTTTAGAGTGAAGTATGATGATGATGATCTTTCA 571
 1292 AGGCCCTTGTGATATGAGCACTGTCATTCATGTCGCTTTGGCATTCATCTTGCCA 1351
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 570 AGGCATTTATGACATGACGCGTTCATTCATCTGCGCTTTGATACATATCCCTA 511
 1352 CAGTAGAGTAGAGAGAGAGAGAAATATCAACGCTTAAGACAGCTCTTGAAGGCA 1411
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 510 CGAAGATTCATGAAAGAGAGAAAGAACTCCAAATATTTATGAAAGCTGCTGATGCT 451
 1412 CTAGAGGCGCCAGAAACGATATATCATATCTATGTCAGAGAAAGGAGGTTGAGATGG 1471

DB 450 CCAGCAAAAGAGAGACATATACACTTATTTGTGAGAGAGAGGCGGTTAAGACAG 391

QY 1472 AGTCCATTACCGAGAGTGTAGCAATTTACCTGATGCTGTGAGGCTGCTGTGTT 1531

DB 390 AGAAGATTTATCTAATTAATGATGATGATGATGATGATGATGATGATGATGATG 331

QY 1532 TGAAGTTCATACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1591

DB 330 ATTAATTTAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 271

QY 1592 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1651

DB 270 GAGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 211

QY 1652 CTCTTATACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1711

DB 210 CGTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 151

QY 1712 CTGTGCAAGTTTGAAGTTCATTTTGTGTAGC 1744

DB 150 CATGTGCAAGCTTGAAGTTCATTTTGTGTATAC 118

RESULT 6
BQ167130/c 701 bp mRNA linear EST 25-APR-2002

LOCUS WHE0946_F09_K182Y Wheat 5-15 DAP spike cDNA library Triticum 391

DEFINITION WHE0946_F09_K18, mRNA sequence.

ACCESSION BQ167130

VERSION BQ167130.1 GI:20311201

KEYWORDS EST.

SOURCE bread wheat.

ORGANISM Triticum aestivum

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticaceae; Triticum.
1 (bases 1 to 701)
Anderson, O.D., Cho, S., Choi, D.W., Close, T.J., Fenton, R.D., Han, P.S., Hsieh, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J., Seaton, C.L. and Tong, J.C.

TITLE The structure and function of the expressed portion of the wheat genomes - 5-15 DAP spike cDNA library

JOURNAL Unpublished (2000)

COMMENT Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818

FEATURES
source Email: oanderson@w.usda.gov
This EST was generated by sequencing from the 3' end of the clone. Sequences have been trimmed to remove vector sequence and low quality sequence with phred score less than 20.
Seq primer: Oligo dt wobble primer (an equal mixture of (T)27A, (T)27G and (T)27C).

Location/Qualifiers
1..701
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE0946_F09_K18"
/clone_1ib="Wheat 5-15 DAP spike cDNA library"
/tissue_type="spike"
/dev_stage="Adult plant"
/lab_host="E. coli SOLR"
/note="Vector: Lambda uni-zap XR, excised phagemid; Site_1: EcoRI; Site_2: XhoI; Plants were grown in the greenhouse. Spikes at 5, 10 and 15 DAP were harvested, total RNA and poly(A) RNA were prepared, a cDNA library was made, and the cDNA clones were in vivo excised to give Bluescript phagemids in the TJ Close lab (Choi, Close, Fenton) at the University of California,

BASE COUNT 218 a 153 c 124 g 206 t

ORIGIN

Query Match 11.7%; Score 257.8; DB 14; Length 701;
Best Local Similarity 68.2%; Pred. No. 1.7e-36;
Matches 358; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 1220 ATCTAGCAGAAATCCCGAGCATATATATCATCATTTGAGGTGAGTATGAGGTCCAT 1279

DB 701 ATATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 642

QY 1280 CTGACCTCTCAAGGCGGCTGTGATATGACAGGTGACATGCTGAGGCTGCTTGA 1339

DB 641 CTGATCTTTTCAAGGAGCTTATGATGATGATGATGATGATGATGATGATGATGATG 582

QY 1340 TTACCTTGCCACAGTGTAGTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1399

DB 581 TACATATCGGTACGAGAGATTTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 522

QY 1400 CTTCAGAGTCTACAGAGGCGCAGAAACCTATATCATATCATGTGAGAGAGAGAG 1459

DB 521 CATGTAGTCTCTCCAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 462

QY 1460 GGTGTGATGAGAGTCCATTTACCTGAGTGTGATGATGATGATGATGATGATGATG 1519

DB 461 GGTTTAG 402

QY 1520 CTGCTGTTGTTTGAAGTTCATCATTTGATGATGATGATGATGATGATGATGATGAT 1579

DB 401 CTGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 342

QY 1580 CTGAGTTATAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1639

DB 341 CATCAAGTATTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 282

QY 1640 CACAGACATCAAGGCTTATACCTGATGATGATGATGATGATGATGATGATGATGAT 1699

DB 281 CAGAGACATCAAGGCTTATACCTGATGATGATGATGATGATGATGATGATGATGAT 222

QY 1700 TGGAGACATCAAGGCTTATACCTGATGATGATGATGATGATGATGATGATGATGAT 1744

DB 221 TTAAGGCGCATCATGTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 177

RESULT 7
BE412601 856 bp mRNA linear EST 24-JUL-2000

LOCUS MCG002_B07R990625 ITTC MCG Barley leaf/Culm Library Hordeum vulgare 222

DEFINITION MCG002_B07R990625, mRNA sequence.

ACCESSION BE412601

VERSION BE412601.1 GI:9410447

KEYWORDS EST.

SOURCE Hordeum vulgare.

ORGANISM Hordeum vulgare

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticaceae; Hordeum.
1 (bases 1 to 856)
Anderson, O.A., Appels, R., Bailey, P., Blake, T., Close, T., Cloutier, S., Dubcovsky, J., Feuillet, C., Gale, M., Graner, A., Gustafson, P., Herrmann, R.G., Holton, T., Jacquemin, J.M., Jia, J., Joudrier, P., Langridge, P., Lazo, G.R., Lin, J.J., McQuilte, P., Ogihara, Y., Pecchioni, N., Qualset, C., Schuch, W., Selvaraj, G., Shariflou, M., Sorrells, M., Warburton, M. and Wenzel, G.

TITLE International Triticaceae EST Cooperative (ITTC): Production of Expressed Sequence Tags for Species of the Triticaceae

JOURNAL Unpublished (2000)

COMMENT Contact: Graner A
Institute for Plant Genetics & Crop Plant Research
Corrensstr. 3, D-06466 Gatersleben GERMANY

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2003, 01:12:16 ; Search time 117 Seconds
(without alignments)
5761.329 Million cell updates/sec

Title: US-09-913-858A-1

Perfect score: 2198
Sequence: 1 actaactcaacagctgcatt.....aaaaaaaaaaaaaaaaaaaaa 2198

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/prodata1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata1/ina/6A_COMB.seq:*
4: /cgn2_6/prodata1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/prodata1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	209.4	9.5	2069	4	US-09-678-300-3
2	209.4	9.5	2124	4	US-09-678-300-6
3	209.4	9.5	2165	4	US-09-678-300-9
4	209.4	9.5	2232	4	US-09-678-300-12
5	209.4	9.5	2808	4	US-09-678-300-2
6	209.4	9.5	2863	4	US-09-678-300-5
7	209.4	9.5	2904	4	US-09-678-300-8
8	209.4	9.5	2971	4	US-09-678-300-11
9	209.4	9.5	3714	4	US-09-678-300-1
10	44.4	2.0	1639	2	US-08-737-524B-1
11	42.4	1.9	665	4	US-08-998-416-185
12	42.4	1.9	662	4	US-08-998-416-137
13	42.4	1.9	701	4	US-08-998-416-701
14	42.4	1.9	711	4	US-08-998-416-786
15	42.4	1.9	724	4	US-08-998-416-683
16	42.4	1.9	732	4	US-08-998-416-1036
17	42.4	1.9	828	4	US-08-998-416-538
18	42.4	1.9	834	4	US-08-998-416-305
19	42.4	1.9	863	4	US-08-998-416-191
20	42.4	1.9	860	4	US-08-998-416-287
21	40.6	1.8	7218	1	US-08-232-463-14
22	40.4	1.8	8154	4	US-08-998-416-534
23	39.4	1.8	3573	4	US-09-353-585-4
24	38.6	1.8	6124	4	US-08-213-4198-3
25	38.2	1.7	7218	1	US-08-232-463-14
26	38.2	1.7	1454	1	US-08-220-958-3
27	37.8	1.7	723	2	US-08-618-911-1

28	37.8	1.7	767	4	US-08-998-416-472	Sequence 472, App
29	37.8	1.7	827	4	US-08-998-416-535	Sequence 535, App
30	37.8	1.7	1578	4	US-09-416-050A-1	Sequence 1, Appl1
31	37.8	1.7	1578	4	US-09-664-800-1	Sequence 1, Appl1
32	37.8	1.7	1578	4	US-09-665-309-1	Sequence 1, Appl1
33	37.8	1.7	1578	4	US-09-661-569-1	Sequence 1, Appl1
34	37.8	1.7	4673	1	US-07-638-431-1	Sequence 1, Appl1
35	37.8	1.7	4673	5	PCT-US92-00018-1	Sequence 1, Appl1
36	37.6	1.7	1371	2	US-08-428-713-1	Sequence 1, Appl1
37	37.6	1.7	1371	3	US-08-904-179-1	Sequence 1, Appl1
38	37.4	1.7	87350	3	US-08-781-891-79	Sequence 79, Appl1
39	37.4	1.7	87343	4	US-09-791-211-3	Sequence 399, App
40	37.2	1.7	629	4	US-09-385-982-389	Sequence 11, Appl1
41	36.8	1.7	1058	4	US-09-452-239-11	Sequence 1, Appl1
42	36.8	1.7	1353	4	US-09-390-131-1	Sequence 3, Appl1
43	36.2	1.6	2073	4	US-09-173-300-3	Sequence 1, Appl1
44	36.2	1.6	2002	4	US-09-819-993-1	Sequence 1, Appl1
45	35.8	1.6	1026	4	US-09-116-498-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-09-678-300-3
; Sequence 3, Application US/09678300
; Patent No. 6420548
; GENERAL INFORMATION:
; APPLICANT: VOZINA, Louis-Philippe
; APPLICANT: D'Aoust, Marc-Andr
; TITLE OF INVENTION: MEDICAGO INC.
; TITLE OF INVENTION: METHOD FOR REGULATING TRANSCRIPTION OF
; FILE REFERENCE: 14149-3"PCT"
; CURRENT APPLICATION NUMBER: US/09/678, 300
; CURRENT FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US 60/157,133
; PRIOR FILING DATE: 1999-10-04
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2069
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequences to be used as promoter for regulating
; OTHER INFORMATION: expression of foreign genes
US-09-678-300-3

Query Match 9.5%; Score 209.4; DB 4; Length 2069;
Best Local Similarity 66.5%; Pred. No. 2.9e-51;
Matches 341; Conservative 0; Mismatches 151; Indels 21; Gaps 2;

QY	1486	AGGCTGACCAATTTACTGTAATGCTGTGAAGCGCTGTTGTTGAACTACATCC	1545
DB	170	AGATCTACCAATTTACTGTAATGCTGTGAAGCGCTGTTGTTGAACTACATCC	229
QY	1546	CTGATCTGTCGCTGTGTAAGCAAGCCTGTAAGTTAAAGGGGGAGTCT	1605
DB	230	CTGATCTGTCGCTGTGTAAGCAAGCCTGTAAGTTAAAGGGGGAGTCT	289
QY	1606	TTAAACCTCAAAATTTATCCCAATTTGCTTGACACAGACAGCAAGCTTTATACCTTC	1665
DB	290	TTGAAAGTTTCAAAATTTATCCCTGCGGCTTGACACAGACAGCAAGCTTTATACCTTC	349
QY	1666	AGCTTCAAGGTGATGCTGATTCAGAGTCACTTGAGAACAAATCTTGCCAAAGTTT	1725
DB	350	CAGTTCAAGGGGATGATTCAGAGTCACTTGAGAACAAATCTTGCCAAAGTTT	409
QY	1726	GAATGATTTTGTGTCATGACATGCGCTAA-----ATGATACCTCTGCTACCTGATTT	1778
DB	410	GAATGATTTTGTGTCATGATGAGTACCTGATTAATTTACATGATGATGACATGATTT	469


```

: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Sequences to be used as promoter for regulating
: OTHER INFORMATION: expression of foreign genes
US-09-678-300-11

```



```
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 937:
SEQUENCE CHARACTERISTICS:
LENGTH: 665 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1581RP
US-08-998-416-937

Query Match
Best Local Similarity 46.9%; Score 42.4; DB 4; Length 665;
Matches 165; Conservative 0; Mismatches 186; Indels 1; Gaps 1;

QY 1845 TTATTATGCTAGTCTTGGCCACACATGATGTTTGTATAGACATCACACTTTA 1904
DB 196 TTGATTTTATCTATTTAACAATAAATTTTAAATGTTATTAATAAATAAATAA 255
QY 1905 ATTTTAACTGTTTCTGTAGAAAGTCAATTCATATTTTATGTTTATGCTCT 1964
DB 256 CTATAGAAATATTTATTAATAAGATTTTAATTTTAAATTTTAAATTTTAAATTTT 314
QY 1965 TATCTGATCTCTAGAGTACAGACTTCTGTATTTGTGAGTGAACCTGAATCTATA 2024
DB 315 TATTAATAAATAGATTTTAAAGTTTATTAATAAGTATATTTTAAATTTTAAATA 374
QY 2025 GAAGATCAGATGTTTCAACAGACATTTACTTCACTGTTTGTGATCTCGA 2084
DB 375 AATTATTTATTTTACTTCATGATATATTTAATTTAATTTTAAATTTTAAATTTT 434
QY 2085 GCTTTTATGCTGTGAGACTGCTGCTGTTGAGCAGCTGTTTATGCTTCACTGTTA 2144
DB 435 TTTTATTTAGCTAGTAAATTTCTATTTAATAGTCTACCTTTAATTTGATTTACTAC 494
QY 2145 CTGTCAGTGTCTGTTTGTGACCTCTAATAAAAAAAAAAAAAAAAAAAAAA 2196
DB 495 CTACTAATAATTACCTAATAATATTTATTAAGAAATCTTAATCTAATAA 546

RESULT 13
US-08-998-416-701
Sequence 701, Application US/08998416
GENERAL INFORMATION:
PATENT NO. 6239264
APPLICANT: Philippesen, Peter
APPLICANT: Philippesen, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
APPLICANT: Reibischung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: NO. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NO. 6239264th Carolina
COUNTRY: USA
ZIP: 27709
```

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 701:
SEQUENCE CHARACTERISTICS:
LENGTH: 701 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1462RP
US-08-998-416-701

Query Match
Best Local Similarity 46.9%; Score 42.4; DB 4; Length 701;
Matches 165; Conservative 0; Mismatches 186; Indels 1; Gaps 1;

QY 1845 TTATTATGCTAGTCTTGGCCACATGATGTTTGTATAGACATCACACTTTA 1904
DB 196 TTGATTTTATCTATTTAACAATAAATTTTAAATGTTATTAATAAATAAATAA 255
QY 1905 ATTTTAACTGTTTCTGTAGAAAGTCAATTCATATTTTATGTTTATGCTCT 1964
DB 256 CTATAGAAATATTTATTAATAAGATTTTAATTTTAAATTTTAAATTTTAAATTTT 314
QY 1965 TATCTGATCTCTAGAGTACAGACTTCTGTATTTGTGAGTGAACCTGAATCTATA 2024
DB 315 TATTAATAAATAGATTTTAAAGTTTATTAATAAGTATATTTTAAATTTTAAATA 374
QY 2025 GAAGATCAGATGTTTCACTCAGACACATTTACTTCACTGTTTGTGATCTCGA 2084
DB 375 AATTATTTATTTTACTTCATGATATATTTAATTTAATTTAATTTAATTTAATTTT 434
QY 2085 GCTTTTATGCTGTGAGACTGCTGCTGTTGAGCAGCTGTTTATGCTTCACTGTTA 2144
DB 435 TTTTATTTAGCTAGTAAATTTCTATTTAATAGTCTACCTTTAATTTGATTTACTAC 494
QY 2145 CTGTCAGTGTCTGTTTGTGACCTCTAATAAAAAAAAAAAAAAAAAAAAAA 2196
DB 495 CTACTAATAATTACCTAATAATATTTATTAAGAAATCTTAATCTAATAA 546

RESULT 14
US-08-998-416-786
Sequence 786, Application US/08998416
GENERAL INFORMATION:
PATENT NO. 6239264
APPLICANT: Philippesen, Peter
APPLICANT: Philippesen, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
APPLICANT: Reibischung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
TITLE OF INVENTION: AND USES THEREOF
```

```

:
: NUMBER OF SEQUENCES: 1152
:
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: No. 6239264rtis Corporation
: STREET: 3054 Cornwallis Road
: CITY: Research Triangle Park
: STATE: No. 6239264th Carolina
: COUNTRY: USA
: ZIP: 27709
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
:
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/998,416
: FILING DATE: 24-DEC-1997
: CLASSIFICATION: 435
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: CH 0016/97
: FILING DATE: 31-DEC-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Meigs, J. Timothy
: REGISTRATION NUMBER: 38,241
: REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 919-541-8589
: TELEFAX: 919-541-8587
: INFORMATION FOR SEQ ID NO: 786:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 711 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: ORIGINAL SOURCE:
: ORGANISM: PAG1504UP
:
: US-08-998-416-786
:
: Query Match 1.9%; Score 42.4; DB 4; Length 711;
: Best Local Similarity 46.9%; Pred. No. 0.014;
: Matches 165; Conservative 0; Mismatches 186; Indels 1; Gaps 1;
:
: QY 1845 TTATTTATGCTAGTCTTCTGGCCAACTGATGTTTGTATAGACATCACACTTTA 1904
:   || |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
: Db 197 TTGATTTATATCTATTAACTAAACATTTTAAATGTATTAATAATTAAGAAATTA 256
:
: QY 1905 ATTTTAACTGTTCTGTAGACGCAATCCATTTATGCTTACGTCTCT 1964
:   || |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
: Db 257 CTATATAGATATTTATTAATAGATATTTATTTATTTTAAAT-ATTAATATACCATTTT 315
:
: QY 1965 TATCTGATCATCTGAAGTCACAGTCTTGTATTTGTAGTGAAGAACTGAATCTAATA 2024
:   || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
: Db 316 TATTAATTAATAGATTTATTAAGTTATTAATTAAGATATATTAATTTATTTATTA 375
:
: QY 2025 GAAGATCAGATGTTTCACTCAAGACACATTTACTTCATGTTGTTTGTATGATCTCGA 2084
:   || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
: Db 376 AATTATTTAATTTACTTCATGATATATATTAATTAATTAATGATCTTCATTAATATTTA 435
:
: QY 2085 GCTTTTATGCTGGAAGTCGCCCTGCTGTTGAGACCTGTTATGCTTACGTGTTA 2144
:   || |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
: Db 436 TTTTATATAGCTAGTATATTTCTATTTAATAGCTTACCTTTATTTAGATTTACTAC 495
:
: QY 2145 CTGTCCAGTGTATCGTTTGTGACCTCTAATAAAAAAAAAAAAAAAAAAAAAA 2196
:   || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
: Db 496 CTACTAATATTTACCTAATATATATTTATTAAGATATCTTAATCTAATAA 547
:
: RESULT 15
: US-08-998-416-683
: ; Sequence 683, Application US/08998416
: ; Patent No. 6239264
: ; GENERAL INFORMATION:
: ; APPLICANT: Philippsen, Peter

```

```

:
: APPLICANT: Pohlmann, Rainer
: APPLICANT: Steiner, Sabine
: APPLICANT: Mohr, Christine
: APPLICANT: Wendland, Jurgen
: APPLICANT: Knechtle, Philipp
: APPLICANT: Reibschung, Corinne
: TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
: NUMBER OF SEQUENCES: 1152
:
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: No. 6239264rtis Corporation
: STREET: 3054 Cornwallis Road
: CITY: Research Triangle Park
: STATE: No. 6239264th Carolina
: COUNTRY: USA
: ZIP: 27709
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
:
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/998,416
: FILING DATE: 24-DEC-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: CH 0016/97
: FILING DATE: 31-DEC-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Meigs, J. Timothy
: REGISTRATION NUMBER: 38,241
: REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 919-541-8589
: TELEFAX: 919-541-8587
: INFORMATION FOR SEQ ID NO: 683:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 724 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: ORIGINAL SOURCE:
: ORGANISM: PAG1453RP
:
: US-08-998-416-683
:
: Query Match 1.9%; Score 42.4; DB 4; Length 724;
: Best Local Similarity 46.9%; Pred. No. 0.014;
: Matches 165; Conservative 0; Mismatches 186; Indels 1; Gaps 1;
:
: QY 1845 TTATTTATGCTAGTCTTCTGGCCAACTGATGTTTGTATAGACATCACACTTTA 1904
:   || |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
: Db 196 TTGATTTATATCTATTAACTAAACATTTTAAATGTATTAATAATTAAGAAATTA 255
:
: QY 1905 ATTTTAACTGTTCTGTAGACGCAATCCATTTTATGCTTACGTCTCT 1964
:   || |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
: Db 256 CTATTAAGATATTTATTAATAGATATTTATTTATTTTAAAT-ATTAATATACCATTTT 314
:
: QY 1965 TATCTGATCATCTGAAGTCACAGTCTTGTATTTGTAGTGAAGAACTGAATCTAATA 2024
:   || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
: Db 315 TATTAATTAATTAATAGATTTATTAAGTTATTAATTAAGATATTAATTTATTTAATA 374
:
: QY 2025 GAAGATCAGATGTTTCACTCAAGACACATTTACTTCATGTTGTTTGTATGATCTCGA 2084
:   || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
: Db 375 AATTATTTAATTTACTTCATGATATATATTTATTTAATGATCTTTCAATTAATTTA 434
:
: QY 2085 GCTTTTATGCTGGAAGTCGCCCTGCTGTTGAGACCTGTTATGCTTACGTGTTA 2144
:   || |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
: Db 435 TTTTATATAGCTAGTATATTTCTATTTAATAGCTTACCTTTATTTAGATTTACTAC 494
:
: QY 2145 CTGTCCAGTGTATCGTTTGTGACCTCTAATAAAAAAAAAAAAAAAAAAAAAA 2196
:   || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
: Db 495 CTACTAATATTTACCTAATATATATTTATTAAGATATCTTAATCTAATAA 546

```

Mon Apr 28 09:35:59 2003

Search completed: April 28, 2003, 03:45:53
Job time : 157 secs

us-09-913-858a-1.rni

Query Match	2.2%	Score 49.2	DB 9	Length 477
Best Local Similarity	49.2%	Pred. No. 0.0017		
Matches 129	Conservative 0	Mismatches 133	Indels 0	Gaps 0
OY	1007	TGGAAGCCCTGAGCAGCTACAAATTTACGCTTAGCGCTTTGAAAATTGCAATGAGGAGATT	1066	
Db	75	TGCGTTTCTTTCGCCGATATAAGTTCAGCTGGCCCTGGAANAATCCATCTGTACAGGCT	134	

```

RESULT 5
US-09-925-301-133
: Sequence 133, Application US/09925301
: Patent No. US20020052508A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA106
: CURRENT APPLICATION NUMBER: US/09/925,301
: CURRENT FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: PCT/US00/05882
: PRIOR FILING DATE: 2000-03-08
: PRIOR APPLICATION NUMBER: 60/124,270
: PRIOR FILING DATE: 1999-03-12
: NUMBER OF SEQ ID NOS: 1694
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 133

```

```
; LENGTH: 1373
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (403)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-133
```

Query Match 1.9%: Score 41; DB 10; Length 1373;

Best Local Similarity 55.2%; Pred. No. 0.72; Mismatches 65; Indels 0; Gaps 0;

```
Matches 80; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 2054 TTATTACTCATGTTGTTGATGATCTCGAGCTTTTGTAGTGTGGAAGTGCCTGT 2113
    || || || || || || || || || || || || || || || || || || || ||
Db 1219 TTGTATTATTAATTTGCTTTGGTATGATTAATGATGATCATGCTTCCCAAT 1278

QY 2114 GCTTGAAGCACTGTTATGCTTCAGTGTACTGTCCAGTGTGTTAGCTTTGACCTCT 2173
    || || || || || || || || || || || || || || || || || || || ||
Db 1279 TGTGGGTGCTCTGTGGCGCATCAATAAAGCCGCTCTGATTTATTTTCAAAAAA 1338

QY 2174 AAAAAAAAAAAAAAAAAAAAAA 2198
Db 1339 AAAAAAAAAAAAAAAAAAAAAA 1363
```

RESULT 6

```
US-09-925-302-245
; Sequence 245, Application US/09925302
; Patent No.: US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 245
; LENGTH: 4065
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-302-245
```

Query Match 1.9%: Score 41; DB 10; Length 4065;

Best Local Similarity 50.8%; Pred. No. 1.5; Mismatches 95; Indels 0; Gaps 0;

```
Matches 98; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 2006 TGAAGAACTGAATCTATAGAGATCAGATGTTTCACTCAAGACATTTACTTCAT 2055
    || || || || || || || || || || || || || || || || || || || ||
Db 3853 TCAACCATCTAATCTCTAAGATTTTGTCTCTCTTCCACAAACAGTGTATAACAG 3912

QY 2066 GTTGTATGATGATCTGAGCTTTTGTAGTCTGGAACCTCCCTGTTGAGCACC 2125
    || || || || || || || || || || || || || || || || || || || ||
Db 3913 AAATACAGTACGCTGTTTCTCTCTGTGTGTAAGTAATGATGATGATGATGACT 3972

QY 2126 TGTATTGCTTCACTGTTACTGTCCAGTGTGTTATCGTTTGAACCTTAACAAAAA 2185
    || || || || || || || || || || || || || || || || || || || ||
Db 3973 TGTATTGATCAATTAACACTAAGAAATTAACATTCCTTTAAAAA 4032

QY 2186 AAAAAAAAAAAAAA 2198
Db 4033 AAAAAAAAAAAAAA 4045
```

RESULT 7

```
US-10-123-155-412
; Sequence 412, Application US/10123155
; Publication No. US20030068794A1
```

GENERAL INFORMATION:

```
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Desnoyers, Luc
; APPLICANT: Flivaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Matanabe, Colin K
; APPLICANT: Zhang, Zemin
```

```
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
```

```
; PRIOR APPLICATION REMOVED - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 412
```

```
; LENGTH: 1184
; TYPE: PRT
```

```
; ORGANISM: Homo Sapien
US-10-123-155-412
```

Query Match 1.8%: Score 40.2; DB 9; Length 1184;

Best Local Similarity 8.5%; Pred. No. 1.1; Mismatches 419; Indels 0; Gaps 0;

```
Matches 56; Conservative 186; Mismatches 419; Indels 0; Gaps 0;

QY 882 TTTCAATTCATGTTGGTGTCTGCAAAATTTCCGTTGCAAGCTCTTGAGCCCTTGAAA 941
    || || || || || || || || || || || || || || || || || || || ||
Db 249 TPKLLTQDSGRRRIPCLCPDGKSLIKITVVKFAPVILTPKXSLKATIKAEFVRRET 308

QY 942 ATCAACATCAAAATTTGATTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1001
    || || || || || || || || || || || || || || || || || || || ||
Db 309 PYWVWNPETKARRAGQSVSLCKKATKRPDPKPYFWYHNDTLDDSLVKKHESLVLRLQ 368

QY 1002 CAAAGTGAAGCCCTGAAGCACTCAAAATTTAGCTTACGCTTTGAAATTTGAATGAGA 1061
    || || || || || || || || || || || || || || || || || || || ||
Db 369 HQAGEYFCKRQSDAGAVKSKAQLIVTASDTPCNPVESILIRLPHCFONATNSFIYD 428

QY 1062 AGATTATGTAAGTGAAGAAATTTCTTCAATCCCTGTTGGAACGTGCTGCTGTTGT 1121
    || || || || || || || || || || || || || || || || || || || ||
Db 429 VGRCPVKTCAGQDONGIICRDVONCCGISKTEREIOCGSVTLPTFKAKKSCORCTET 488

QY 1122 TGTGCTCCCAATATATGAGACTTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 1181
    || || || || || || || || || || || || || || || || || || || ||
Db 489 RSIVGRVSAADNGEPMRFGHVYMGNSRVSMYGRKFTTLHVPDTERLVLFVDRLOKE 548

QY 1182 GATGAGAGATGTTAGTCTGTGTCGAAGACCATGATATCTACAGCAAAATCCCGAAGC 1241
    || || || || || || || || || || || || || || || || || || || ||
Db 549 VNTTKVLPFNKGSVAFHEIKMLRKEPTILEAMETNIILPGEVVGDEPMALPEIPSRF 608

QY 1242 ATATATCAATCATTTAGCTGGAAGTATGAGGTTCCATCTGCACTGCTCAAGGCCCTTGT 1301
    || || || || || || || || || || || || || || || || || || || ||
Db 609 YRONGEPIYIGKVASVFLDPRNISTATAQTDLNFIDEEDTFLPLRTYGMFSVDFRDEV 668

QY 1302 GGATATGCGACCTGCTGATTCATCGTCCGCTTTTGCATTCATCTGGCCACAGGAGTAG 1361
    || || || || || || || || || || || || || || || || || || || ||
Db 669 TSEPLNAGKVVHLDSITOVKMEHISTYKILWSLMPDGTGLWEEBDFENORRNKRREDRT 728

QY 1362 AGAGAAGAGAAGAAATATCAACGCTTAAGAGACGCTCTGCAAGTGCACATAGAGGCC 1421
    || || || || || || || || || || || || || || || || || || || ||
Db 729 FLVGNLEIRERLFLNDVPESRCFVKYRAISRERFLPSBEOIGVVISVILMERPTGLS 788

QY 1422 AGAAACGTAATATCATATCTATGTCAGAGAAAGGGAAGTTTGAGATGAGATGCATTTA 1481
    || || || || || || || || || || || || || || || || || || || ||
```

Db 789 NFRANGREDVITGPNGACVPAPFCDDSDPAYSAVILASLAGEELQAVESSEKFNPNMIG 848
QY 1482 CCTGAGCTGTAGCAATTAACCTGATGCTGTGAAGGCTGCTGTTGTAAGTTCAC 1541
Db 849 VPQPLMLNLRTRDHPKRYKTAFOISMARPNSENGPIYAFENLRACEAPPS 908
QY 1542 A 1542
Db 909 A 909

RESULT 8
US-09-731-872-43
; Sequence 43, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bouquellet, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/09/731.872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 43
; LENGTH: 1942
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 334..1551
; NAME/KEY: sig_peptide
; LOCATION: 334..426
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4.0554926521937
; OTHER INFORMATION: seq TVFLVLTQLDLT/VE
US-09-731-872-43

Query Match 1.8%; Score 40.2; DB 10; Length 1942;
Best Local Similarity 47.1%; Pred. No. 1.5;
Matches 123; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

QY 1021 CACTACAAATTAAGCTTGCCTTTGAAATTCGATGAGGAAGATTATGTAACTGAAAA 1080
Db 1276 CAGTATTAAGTTATTCCTTGAAGTTCAGTTGTGATGACTACATCAGTAGAAG 1335
QY 1081 TTCTTCCAAATCCCTGTGTGAGAACTGCTGTGTGTGTGTGCTCCAAATTTAG 1140
Db 1336 TTCTGAGGCCACGTAACACTGGGGTAGTCCCTGTATTTACGATATCCCCAGCATCA 1395
QY 1141 GACTTGTCTCTTCTCTGTTCAATTTTACATATTTAAGAGATAGAGATGTGAGTCT 1200
Db 1396 GACTGGCTTCCAGTAACAAAGTGTCTTGTATCAGAAATTTTTCACCCCGGAA 1455
QY 1201 GTTGCAAGACCATGAGATATCTAGCAGAAAAATCCCGAAGCATATTAATCAATTCAGG 1260
Db 1456 CTGGCAAGTTTACATCAGACGACTGATTTGTATGACAGATTTGTATGAGGCTATGTAGAA 1515
QY 1261 TGAAGTATGAGGTCATCT 1281
Db 1516 TGGAACTGGAAGGGTAGATCT 1536

RESULT 9
US-09-764-853-180
; Sequence 180, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P1206
; CURRENT APPLICATION NUMBER: US/09/764.853
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: Patentlin Ver. 2.0
; SEQ ID NO 180
; LENGTH: 1446
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1446)
; OTHER INFORMATION:
US-09-764-853-180

Query Match 1.8%; Score 39; DB 10; Length 1446;
Best Local Similarity 49.3%; Pred. No. 2.7;
Matches 99; Conservative 1; Mismatches 101; Indels 0; Gaps 0;

QY 1998 ATTGTGAGTGAAGAACTGAATCTAATAGAGATCAGATGTTTACTCAGACACATTTAT 2057
Db 1238 AATAGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1297
QY 2058 TACTCATGTTGTTTGTGATGATCTGACCTTTTGTAGTGTGGAAGTGTCCGTGTT 2117
Db 1298 TTCAATGGAATTAACCTACTGATCTTGTCTGTGCAAGTACAGACTGACCTTCTTT 1357
QY 2118 TGAGCACCTGTTATTCCTTCAGTGTCTGACGTGTGATGCTTTTGTGACTGTAAA 2177
Db 1358 CAAGTACCTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1417
QY 2178 AAAAAAAAAAAAAAAAAAAAAA 2198
Db 1418 AAAAAAAAAAAAAAAAAAAAAA 1438

RESULT 10
US-09-764-853-370/c
; Sequence 370, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P1206
; CURRENT APPLICATION NUMBER: US/09/764.853
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: Patentlin Ver. 2.0
; SEQ ID NO 370
; LENGTH: 1446
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION:
US-09-764-853-370

Query Match 1.8%; Score 39; DB 10; Length 1446;
Best Local Similarity 49.3%; Pred. No. 2.7;
Matches 99; Conservative 1; Mismatches 101; Indels 0; Gaps 0;

QY 1998 ATTGTGAGTGAAGAACTGAATCTAATAGAGATCAGATGTTTACTCAGACACATTTAT 2057
Db 209 AATAGAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 150
QY 2058 TACTTCATGTTGTTTGTATGATCTGACCTTTTGTAGTGTGGAAGTGTCCGTGTT 2117
Db 149 TTCAATGGAATTAACCTACTGATCTTGTCTGTGCAAGTAAACAGACTAGAACCTTCTTT 90

APPLICANT: Slader

ORGANISM: *Arabidopsis thaliana*

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2003, 00:45:31 ; Search time,5603 Seconds

(without alignments)
11416.735 Million cell updates/sec

Title: US-09-913-858a-1

Perfect score: 2198
Sequence: 1 actactcaacacgtcgtcatt.....aaaaaaaaaaaaaaaaa 2198

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_cm:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vit:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Prod. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2198	100.0	2198	6 AX033383	AX033383 Sequence
2	2191.6	99.7	2198	6 VRA18529	V18529 Vigna radia
3	690	31.4	1508	8 ATH404860	ATH404860 Arabidops
4	690	31.4	1729	8 ATH345084	AJ345084 Arabidops
5	678.8	30.9	1542	8 ATH345085	AJ345085 Arabidops
6	678.8	30.9	1938	8 AY054522	AY054522 Arabidops
7	664.8	30.2	1834	8 ATH404861	ATH404861 Arabidops
8	534	24.3	658	8 MTR416755	AJ416755 Medicago
9	349	15.9	787	8 AF277228	AF277228 Arabidops
10	346	15.7	588	8 AF277229	AF277229 Arabidops
11	209.4	9.5	2069	6 AX105537	AX105537 Sequence
12	209.4	9.5	2124	6 AX105540	AX105540 Sequence
13	209.4	9.5	2165	6 AX105543	AX105543 Sequence
14	209.4	9.5	2232	6 AX105546	AX105546 Sequence
15	209.4	9.5	2808	6 AX105536	AX105536 Sequence
16	209.4	9.5	2863	6 AX105539	AX105539 Sequence
17	209.4	9.5	2904	6 AX105542	AX105542 Sequence
18	209.4	9.5	2971	6 AX105545	AX105545 Sequence
19	209.4	9.5	3714	6 AX105535	AX105535 Sequence
20	178.8	8.1	104679	8 AC011807	AC011807 Arabidops
21	171	7.8	81875	8 AP000419	AP000419 Arabidops
22	155	7.1	89134	2 AP004162	AP004162 Oryza sat
23	155	7.1	141717	2 AP004457	AP004457 Oryza sat
24	105	4.8	105	6 AX033385	AX033385 Sequence
25	60.2	2.7	1449	8 BVU315848	AJ315848 Beta vulg
26	60	2.7	1630	8 LES313193	AJ313193 Lycopersi
27	54.6	2.5	177674	2 AC127920	AC127920 Rattus no
28	54.6	2.5	190304	2 AC114439	AC114439 Rattus no
29	53.4	2.4	97683	2 AC116548	AC116548 Dictyoste
30	53.4	2.4	234488	2 AC109698	AC109698 Rattus no
31	807	2.4	807	8 MTR416757	AJ416757 Medicago
32	51.8	2.4	188464	9 AC096757	AC096757 Homo sapi
33	50.6	2.3	122160	2 CNS0787	AL731761 Oryza sat
34	50	2.3	1149	3 CEL505020	AJ505020 Caenorhab
35	50	2.3	39370	3 AC006674	AC006674 Caenorhab
36	50	2.3	180664	2 AC006706	AC006706 Caenorhab
37	50	2.3	183800	2 AC123118	AC123118 Rattus no
38	49.8	2.3	241109	2 AC121599	AC121599 Mus muscu
39	49.6	2.3	1128	10 AF345993	AF345993 Rattus no
40	49.4	2.2	1182	8 AY026941	AY026941 Arabidops
41	49.4	2.2	1209	8 ATH404862	ATH404862 Arabidops
42	49.4	2.2	56956	2 AC117081	AC117081 Dictyoste
43	49.4	2.2	80413	8 AC021665	AC021665 Arabidops
44	49.2	2.2	1479	6 AX356085	AX356085 Sequence
45	49.2	2.2	2408	9 AK095482	AK095482 Homo sapi

ALIGNMENTS

RESULT 1	AX033383	2198 bp	DNA	Linear	PAT 21-SEP-2000
LOCUS	AX033383				
DEFINITION	Sequence 1 from Patent WO0049153.				
ACCESSION	AX033383				
VERSION	AX033383.1 GI:10280157				
KEYWORDS					
SOURCE	unidentified.				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 2198)				
AUTHORS	Altman, F.				
TITLE	Fucosyl transferase gene				
JOURNAL	Patent: WO 0049153-A 1 24-AUG-2000;				
	ALTMANN FRIEDRICH (AT)				

Db 1981 AGTCACAGCTTCTGTATATGAGTGAAGAACTGAATCTAATAGAGATCATGATGTTT 2040
 QY 2041 CACTCAAGACACATTTATTTATCTGATGTTTGTGATCATCGAGCTTTTATGATGCTG 2100
 Db 2041 CACTCAAGACACATTTATTTATCTGATGTTTGTGATCATCGAGCTTTTATGATGCTG 2100
 QY 2101 GAATGTCCTGCTGTTGAGACCTGTTATGCTTCACTGATTCAGTGGTATC 2160
 Db 2101 GAATGTCCTGCTGTTGAGACCTGTTATGCTTCACTGATTCAGTGGTATC 2160
 QY 2161 GTTTTGTACCTCTAATAAAAAAAAAAAAAAAAAAAAAA 2198
 Db 2161 GTTTTGTACCTCTAATAAAAAAAAAAAAAAAAAAAAAA 2198
 RESULT 2
 VRA18529 2198 bp mRNA linear PLN 01-NOV-2001
 LOCUS VRA18529
 DEFINITION Vigna radiata mRNA for alpha-L-3-fucosyltransferase (Fuct c3).
 ACCESSION Y18529
 VERSION Y18529.1 GI:5702038
 KEYWORDS alpha-L-3-fucosyltransferase; fuct gene; GDP-L-Fuc.
 SOURCE mung bean.
 ORGANISM Vigna radiata var. radiata
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Vigna.
 REFERENCE 1
 AUTHORS Leiter, H., Mucha, J., Staudacher, E., Grimm, R., Glossl, J. and
 Altman, F.
 TITLE Purification, cDNA cloning, and expression of GDP-L-Fuc:Asn-Linked
 JOURNAL G1NAc alpha-L-3-fucosyltransferase from mung beans
 MEDLINE J. Biol. Chem. 274 (31), 21830-21839 (1999)
 PUBMED 99348317
 REFERENCE 2 (bases 1 to 2198)
 AUTHORS Mucha, J.
 TITLE Direct Submission
 JOURNAL Submitted (10-DEC-1998) J. Mucha, Institut fuer Chemie, Universitaet
 fuer Bodenkultur Wien, Muthgasse 18, 1190 Wien, Austria
 FEATURES
 source
 1. 2198
 /organism="Vigna radiata var. radiata"
 /db_xref="taxon:3916"
 /issue_type="sprouts"
 211..1743
 /gene="fuct"
 211..1743
 /gene="fuct"
 /codon_start=1
 /product="Fuct c3 protein"
 /protein_id="CAB52254.1"
 /db_xref="GI:5702038"
 /translating="MMGLTLNLSRSDGADQDLSLVLAPGNPKRWSMLPLVAL
 VVAIAELVLRDLMAKNAAMVDSLADFEYSRAVREDDDLGIVASDRSESYSCEE
 WLEREDAVTYSRDESKSEPIFVSGADDEKSCVSGCKRFGSGDRPDAAPRLPSGTA
 SIIRSMESAEYVARNINAMARRGYNIVHTSSSDVPEVFESMAEDMAPVOPKTE
 AALAAFLISGCGARNRFLQALELEKSNIKIDISYGGCHRRDRVNVKLFKFSL
 AFENSNEDYVTEKFOSLVAGTPVYVAGAPNIDAPSPGSLIHLKEIDVESVAKT
 MRYLANPEAYNOSLRMKYEGSPDSFKALVDMAAVHSCGLCHLIVASREKENNS
 LKRPPCKCTGPEYVHYHYVERGREGMESIYRSSNLTLNAYKAAVLFKTSNLVP
 VKTERPEVIRGSSALKLYKIYPIGLQRALTYTFKGDADRSHLENNPYAKFEVI
 FV"
 BASE COUNT 595 a 412 c 516 g 675 t
 ORIGIN
 Query Match 99.7% Score 2191.6; DB 8; Length 2198;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2194; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 ACTAATCAAAACGCTGATTTCTTTTCTTCAGGAGAACCATCCACCATAACAACAA 60

Db 1 ACTAATCAAAACGCTGATTTCTTTTCTTCAGGAGAACCATCCACCATAACAACAA 60
 QY 61 AAAAAACAAGACAGCTGCTTTTATTCGTTCTTTTCTTAACAAGACCCCCA 120
 Db 61 AAAAAACAAGACAGCTGCTTTTATTCGTTCTTTTCTTAACAAGACCCCCA 120
 QY 121 TCATGATGCTGCTCAATACGCCAAATTTCCATTTCCCTTGATTTAGTTATTT 180
 Db 121 TCATGATGCTGCTCAATACGCCAAATTTCCATTTCCCTTGATTTAGTTATTT 180
 QY 181 TCGGAATTTGGACATGAGGGGCGCAATTAATGATGAGGTGTTGACGAATCTTGAGGC 240
 Db 181 TCGGAATTTGGACATGAGGGGCGCAATTAATGATGAGGTGTTGACGAATCTTGAGGC 240
 QY 241 TCGAGACAGATGCTGCGCCAAACAGACAGCTTACCCTTTGGCTCGGAGGACACCA 300
 Db 241 TCGAGACAGATGCTGCGCCAAACAGACAGCTTACCCTTTGGCTCGGAGGACACCA 300
 QY 301 AAGAGAAATGAGACCAATCTAATGCTCTGTTGTTGCTTGGCTTGTGATCGCGAGATC 360
 Db 301 AAGAGAAATGAGACCAATCTAATGCTCTGTTGTTGCTTGGCTTGTGATCGCGAGATC 360
 QY 361 GCGTTCTGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
 Db 361 GCGTTCTGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
 QY 421 TTTCTTACCGCTCTCGAGCGGTGTTGAAGTACCATTTGGGGTGGTGGTGGTGGTGGTGG 480
 Db 421 TTTCTTACCGCTCTCGAGCGGTGTTGAAGTACCATTTGGGGTGGTGGTGGTGGTGGTGG 480
 QY 481 TCTGATGGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
 Db 481 TCTGATGGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
 QY 541 ACGTATTCGAGGCGCTTTTCCAAAGAGCTATTTTGTGTTGCTGAGCTGATCAGAGTGG 600
 Db 541 ACGTATTCGAGGCGCTTTTCCAAAGAGCTATTTTGTGTTGCTGAGCTGATCAGAGTGG 600
 QY 601 AAGCGGTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
 Db 601 AAGCGGTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
 QY 661 TTTGGTACCTCAACCAAGTGAAGACAGTTCATTCGATCAATGATGATGATGATGATGATG 720
 Db 661 TTTGGTACCTCAACCAAGTGAAGACAGTTCATTCGATCAATGATGATGATGATGATGATG 720
 QY 721 TACTATGCTGAGAACCAATATTGCAATGAGCAAGACGAGGAGATTAACATGCTAATGACA 780
 Db 721 TACTATGCTGAGAACCAATATTGCAATGAGCAAGACGAGGAGATTAACATGCTAATGACA 780
 QY 781 ACCAGCTATCTGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 Db 781 ACCAGCTATCTGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 QY 841 GCACAGTGCAGCGCAAAACAGAGCTGCTTCCAGCTGCTTCCATTTCCCAATTTGGT 900
 Db 841 GCACAGTGCAGCGCAAAACAGAGCTGCTTCCAGCTGCTTCCATTTCCCAATTTGGT 900
 QY 901 GCTGGAATTTCCGTTGCAAGCTCTTGAAGCCCTTGAATAATCAAAATCAAAATTTGAT 960
 Db 901 GCTGGAATTTCCGTTGCAAGCTCTTGAAGCCCTTGAATAATCAAAATCAAAATTTGAT 960
 QY 961 TCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 Db 961 TCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 QY 1021 CACTACAATTTAGCTTACCTTGAATAATTCGATGAGAGATTAATGATGATGATGATGATGATG 1080
 Db 1021 CACTACAATTTAGCTTACCTTGAATAATTCGATGAGAGATTAATGATGATGATGATGATGATG 1080
 QY 1081 TTTCTCAATCCTGTTGCTGAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
 Db 1081 TTTCTCAATCCTGTTGCTGAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140

Db 1081 TTTCCCAATCCCTGTTGCTGGAACGTGCTCCGTGTTGTTGCTGCTCAATATTCAG 1140
 Qy 1141 GACCTTGCCTCTCTCTGTTCAATTTTACATATTTAAAGAGATGAGATGTTGAGTCT 1200
 Db 1141 GACCTTGCCTCTCTCTGTTCAATTTTACATATTTAAAGAGATGAGATGTTGAGTCT 1200
 Qy 1201 GTTGCAAAAGACCATGATATCTAGCAGAAAATCCGGAAGCATATATCAATCATTTGAG 1260
 Db 1201 GTTGCAAAAGACCATGATATCTAGCAGAAAATCCGGAAGCATATATCAATCATTTGAG 1260
 Qy 1261 TGGAGATGAGAGGTCATCTGACCTCTTCAAGGCCCTTGGATATGGCAGCTGTCAT 1320
 Db 1261 TGGAGATGAGAGGTCATCTGACCTCTTCAAGGCCCTTGGATATGGCAGCTGTCAT 1320
 Qy 1321 TCATCTGCTCTCTTGTGATTCATCTTGGCCACAGTGAAGAGAGAGAGAAATATAT 1380
 Db 1321 TCATCTGCTCTCTTGTGATTCATCTTGGCCACAGTGAAGAGAGAGAGAAATATAT 1380
 Qy 1381 CCAAGCCTTAAGAGACGCTCTGCAAGTGCATAGAGGGCCAGAAACGCTATATCATATC 1440
 Db 1381 CCAAGCCTTAAGAGACGCTCTGCAAGTGCATAGAGGGCCAGAAACGCTATATCATATC 1440
 Qy 1441 TATGTCAGAGAAAGGAGGTTGATGATGAGATGATGATGATGATGATGATGATGATG 1500
 Db 1441 TATGTCAGAGAAAGGAGGTTGATGATGAGATGATGATGATGATGATGATGATGATG 1500
 Qy 1501 ACTCTGAATGCTGTGAAGAGCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGCT 1560
 Db 1501 ACTCTGAATGCTGTGAAGAGCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGCT 1560
 Qy 1561 GTATGGAAGACTGGAAGGCTGGAAGTTTAAAGAGGGGAGGCTTTAAACTCTACAAA 1620
 Db 1561 GTATGGAAGACTGGAAGGCTGGAAGTTTAAAGAGGGGAGGCTTTAAACTCTACAAA 1620
 Qy 1621 ATATACCAATTTGGCTTACACAGACAGACAGCTCTTATACCTTACCTTACCTTACCT 1680
 Db 1621 ATATACCAATTTGGCTTACACAGACAGACAGCTCTTATACCTTACCTTACCTTACCT 1680
 Qy 1681 GCTGATTTACAGAGACTGATGAGAAATCCCTGTTGCCAAGTTGGAAGTCAATTTTGTG 1740
 Db 1681 GCTGATTTACAGAGACTGATGAGAAATCCCTGTTGCCAAGTTGGAAGTCAATTTTGTG 1740
 Qy 1741 TAGCATGGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800
 Db 1741 TAGCATGGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800
 Qy 1801 GCTAGATTTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860
 Db 1801 GCTAGATTTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860
 Qy 1861 TCTTGGCCAACTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1920
 Db 1861 TCTTGGCCAACTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1920
 Qy 1921 TGTAGAAATGCAAAATCCATTTTAACTGATGATGATGATGATGATGATGATGATGATG 1980
 Db 1921 TGTAGAAATGCAAAATCCATTTTAACTGATGATGATGATGATGATGATGATGATGATG 1980
 Qy 1981 AGTCAACAGTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040
 Db 1981 AGTCAACAGTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040
 Qy 2041 CACTCAACAGCATATTTACTTCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2100
 Db 2041 CACTCAACAGCATATTTACTTCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2100
 Qy 2101 GAATGCTGCTGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2160
 Db 2101 GAATGCTGCTGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2160
 Qy 2161 GTTTTGAACCTCTAATAAAAAAAAAAAAAAAAAAAAAA 2198
 Db 2161 GTTTTGAACCTCTAATAAAAAAAAAAAAAAAAAAAAAA 2198

RESULT 3
 ATH404860
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 SOURCE
 1. 1506
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 1. 1506
 /gene="FUCTA"
 1. 1506
 /function="glycosyltransferase; fucose transfer"
 /codon_start=1
 /product="alpha1,3-fucosyltransferase"
 /protein_id="CAC38048.1"
 /db_xref="GI:13992482"
 /translation="MGVFSMLRGKIGLTHEBELPYVANGSTSSSSPSSKRRKVFLL
 PICALVILIEIGLCRLDNASLVDTLTHFTSSSDKLVGSGIEKQEMLERVSVTL
 YSRDFPDLPIFISGNSKDFKSCVDCFTSDPKPPAAGLSHQPTLSLIRSMESA
 OYQENNLACARRKGYDIWMTLSISVDPYGVESMAVDYDAPQPTRELAALAFIS
 NCARNFRLALFALMTNKLKISYGGCHPNRGSVYKVALKHYKRSLEAFNTNEED
 YVTEKFFQSILVAGSPVYVCAFNIEERAPSPDSFLHKQMDYKAAVKKKKYIADND
 AYNOTLWKHEDGSDSKRALIDMAVHSSCRCLFVATRLREDEKSPERRRCKCT
 RGSSTVHLVYRERGRPMESIFLKDNLLEALEASVLAKEFMSLREYPIWKREPAS
 LRGGKLRVHGIVPIGILROALYNEFEENSSLSSTHIOHNPCKFEVFEV"
 BASE COUNT 413 a 304 c 344 g 445 t
 ORIGIN
 Query Match 31.4%; Score 690; DB 8; Length 1506;
 Best Local Similarity 72.0%; Pred. No. 7.3e-165;
 Matches 900; Conservative 0; Mismatches 350; Indels 0; Gaps 0;

Qy 493 TCTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 552
 Db 256 TCAGATAGAGAAATGCGAGAGGTGTTAGAGAGGATTCAGTTACTTATTTCTAGA 315
 Qy 553 GCGTTTCCAAAGAGCCATTTTGTCTGAGCTGATCAGAGTGAAGTGGTGTGCG 612
 Db 316 GATTTCACTAAAGATCCGATTTTATCTGCTGATGATGATGATGATGATGATGATGATG 375
 Qy 613 GTTGATGATTAATTTGGTTAGTGGGATGAGAAACCGATGCGGCTTTGGTTACT 672
 Db 376 GTTGATGATTAATTTGGTTAGTGGGATGAGAAACCGATGCGGCTTTGGTTAGT 435
 Qy 673 CAACCAAGTGAACAGCTAGATCTGATGATGATGATGATGATGATGATGATGATGATGATG 732
 Db 436 CATCAACCTGGAACACTGATATATCGTTCCATGGAATGAGACACATATTACCAACAG 495

[illegible]

KEYWORDS	core-alpha1,3fucosyltransferase 2; fucT2 gene.
SOURCE	thale cress
ORGANISM	Arabidopsis thaliana
REFERENCE	Arabidopsis thaliana
AUTHORS	Spiering, A.; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eumariophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
TITLE	1 (bases 1 to 1542)
REFERENCE	Baker, H.
AUTHORS	Molecular cloning and characterization of a plant (Beta vulgaris) alpha1,4-fucosyltransferase with specificity for Lewisia and Lewisid synthesis
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 1542)
AUTHORS	Baker, H.
JOURNAL	Direct Submission
TITLE	Submitted (14-sep-2001) Baker H., Plant Research International, Wageningen University and Research centre, droevendaalsesteeg 1, Netherlands
FEATURES	Location/Qualifiers
source	1..1542
	/organism="Arabidopsis thaliana"
	/variety="colombia"
	/db_xref="taxon:3702"
	/tissue_type="seique"
	1..1542
	/gene="fucT2"
	1..1542
	/gene="fucT2"
	/function="glycosyltransferase involved in N-glycan biosynthesis"
	/codon_start=1
	/product="core-alpha1,3fucosyltransferase 2"
	/protein_id="CAC78980.1"
	/db_xref="GI:15722482"
	/translation="MGVEYNLRGPRAGTHDEPATNGSPSSSSPSIRKLSNLLPICALVIAIFGLRLDKALVDITDFDTPSPISQSPAPSDPKKIGLTFPRDSCHEMLRDSVYTSRDTFKDPIFGISGEKDEPMCMVDCPTFGSSSKRTDPAAGLQKQKGLTSTIRSESAQITPENDLQARRRGDIWMTLSVDPVGRISMAEYDIMSVPKRTERTIAAEIFISGARNEFLQALALELAKTNIKIDSYGCHNRNGKVKVEALRYKFEIATENNEDYVTEKREFPQSLVAGSVPVVGPPIIEFAPASDSFLIKTMEDEVPVAKRMVYLAANPAVYNOTLRMKYEGPDSDFKALVDAVAHSSCRCLIFLATRYRDEESSPNFKRCKCSRGSDIYHVEFEREGFEMESYFLNGKSVYOGALSAVYAKRSLKHEAVWKKRPNELKDKELKTHRIYPLGLTGRQALNFKPENGNSLSHIIIONNCARFEVYEV"
BASE COUNT	423 a 302 c 370 g 447 t
ORIGIN	
Query Match	30.9%; Score 678.8; DB 8; Length 1542;
Best Local Similarity	68.4%; Pred. No. 5.3e-162;
Matches	989; Conservative 0; Mismatches 442; Indels 15; Gaps 3;
OY	301 AAGAGGAATGAGCAATCTAATGCTCTGTTGTTGCCCTTGTCATCGCGAGATC 360
DB	109 AAGCGAAATATTCGAATTTGTACACTGCTGCTGCTGTGATGATGCTGAGATC 168
OY	361 GCGTTCTAGGGAGGTGGATATGCGCAAAAAGCGCGCATGGTACCTCCCTGCTAC 420
DB	169 GCGTTTCTGGCTGCGCTGAT-----AAAGTCGCTTGGTGTGATACGTTGACTGAT 219
OY	421 TTTCTACCGCTCGAGCGGCTGCTTGAAGGTGACGATTTGGGGTGGGTTGGTGCT 480
DB	220 TTCTTCACCACTCCGTCACCTCTCGCAGTCTCCACCGCGAGATCGATCGGAAGAAG 279
OY	481 TCTGATCGGAATTCGATCGTATGTTGTTGAGGAGATGTTGGAGAGAGAGATCTCTC 540
DB	280 ATCGGATTAATTACTAGTAG--AGCTGCGAGAGTGGTGTGATGAGAAGATTCAGTT 336
OY	541 ACGTTATGAGAGGCTTTCCAAAGAGCGCTATTGTTGTTGCGAGCTGATCAGAGAGG 600
DB	337 ACTTACTCTAGAGATTTTACTTAACATCCATTTTATCTCTGCGTGAGTGAAGAGACTTT 396
OY	601 AAGTGTCTGGCTGTGATGTAATTTGGGTTTACTGGGAGTAGAAGCCAGATCCCGCA 660

Db 397 CAATGCTGTTCTGTGATGTTGTACATTTGGAGATAGTTTCAGGAAACACCAGATGCTGCG 456
 Oy 661 TTGGGTTACTCTCAACCAAGTGGAGACAGCTTCTGCGATCAATGGAATAGCAGAA 720
 Db 457 TTGGGTTAGGTGAGAAACCTGGAACCTTTAGATATATACGTTCCATGGAATAGCAGAC 516
 Oy 721 TACTATGCTGAGAAACATATTTGGCATGGCAAGAGGGGATATACATCTATATGACA 780
 Db 517 TATATCCAGAAATATGATCTTGCACAGGACGACGAGGAGGTATGATATAGATGAGACC 576
 Oy 781 ACCAGCTATCTTGGATGTTCTCTGTTGGATATTTTTCATGGCTGAGTATGATGATG 840
 Db 577 ACTAGTCTATCATCAGATGTTCTCTGTTGGATATTTTTCGTTGGGCGGAGTATGATGAT 636
 Oy 841 GCACGATGAGCGGAAACCTGAAAGCTGCTGCTGAGCTGCTTCATTTCCATTTGTTGGT 900
 Db 637 TCTCCGCTACAGCAAAACCTGAGAGAGCTATTCAGAGCTGCTTTTATTTTATATGTTGGT 636
 Oy 901 GCTCGAATTTTCGGTTGAGAGCTCTTGAGGCCCTTGAAATCAAAACATCAAAATTTGAT 960
 Db 697 GCTCGAATTTTCGGTTGAGAGCTCTTGAGGCCCTTGAAATCAAAACATCAAAATTTGAT 756
 Oy 961 TCTATAGGTGTTGTACAGAGAACCTGATGAGAGTGAACAAGTGAAGCCCTGAG 1020
 Db 757 TCTATAGGTGTTGTATGATGGAACCGGAGTGAAGAGTTGACAAGTTGAAGCTTTAG 816
 Oy 1021 CACTACAAATTTAGTTAGCTTGAAGATTCGAATGAGGAGATTTATGTAAGCTGAAGAA 1080
 Db 817 CGATACAAATTTAGTTAGCTTGAAGATTCGAATGAGGAGATTTATGTAAGCTGAAGAA 876
 Oy 1081 TTCTTCCAAATTTAGTTAGCTTGAAGATTCGAATGAGGAGATTTATGTAAGCTGAAGAA 1140
 Db 877 TTCTTCCAAATTTAGTTAGCTTGAAGATTCGAATGAGGAGATTTATGTAAGCTGAAGAA 936
 Oy 1141 GACTTGTCT 1200
 Db 937 GAATTTGCGCT 996
 Oy 1201 GTTCAAGACATGAGATATCTAGCAGAAATCCGAGATATATATCAATCAATTTAGG 1260
 Db 997 GTTCAAGACATGAGATATCTAGCAGAAATCCGAGATATATATCAATCAATTTAGG 1056
 Oy 1261 TGGAGATGAGGCTCATCT 1320
 Db 1057 TGGAGATGAGGCTCATCT 1116
 Oy 1321 TCACTGCGCT 1380
 Db 1117 TCT 1176
 Oy 1381 CCAAGCCTTAAAGAGAGCT 1437
 Db 1177 CCAATTTTCAAGAAAGAGAGCT 1236
 Oy 1438 ATCTATGTCAGAAAGAGGAGGTTGAGATGAGTCCATTTACTGAGGCTCTAGCAAT 1497
 Db 1237 GTTTTGTGAGAAAGAGGCGCTTTGAAATGAAATCACTCTTTTGAAGGGTAAAGT 1296
 Oy 1498 TTACTCTGATGCTCTGAAAGGCTGCTGTTTGAAGTTCACATCCCTGAAATCTTGTG 1557
 Db 1297 GTACTCTGAGAACTCTAGATCTGCAATCTCTGCCAAGTCTTTTAAACATGAG 1356
 Oy 1558 CCTGTATGAGAACTGAAAGGCTGAAAGTATTAAGAGGGGAGTCTTAAACCTTAC 1617
 Db 1357 CGAGTGTGAGAAAGAGAGAGGCTGAAACTTAAAGAGAGAGAGAGTCTTAAACATGAT 1416
 Oy 1618 AAAATATACCAATTTGGCTTGAACAAGAGCAAGCTCTTTATACCTTCAAGCTTCAAGAT 1677
 Db 1417 CGATTTTACCGCTTGCCCTTGAACAAGAGCAAGCTTGTAACTCAATTCGAGGAG 1476
 Oy 1678 GAGGCTGATTTTGAAGAGTCTCTGAGAGAAATCTCTGCTCAAGTTTGAAGTATTTT 1737
 Db 1477 AATTGAGTCTTAAGTGTACATTTCAAAACACCTTGTCTAAATTTGAGGTTGTCTTC 1536

Oy 1738 GTGTAG 1743
 Db 1537 GTCTAG 1542

RESULT 6
 AY054522
 LOCUS
 DEFINITION
 Arabidopsis thaliana putative fucosyltransferase (At1g49710;
 F14J22.8) mRNA, complete cds.
 ACCESSION
 AY054522
 VERSION
 AY054522.1
 KEYWORDS
 FLI CDNA.
 SOURCE
 Arabidopsis thaliana.
 ORGANISM
 Arabidopsis thaliana.

REFERENCE
 AUTHORS
 Southwick, A., Karlin-Neumann, G., Nguyen, M., Lam, B., Miranda, M.,
 Cheuk, R., Chung, M.K., Hayashizaki, T., Ishida, J., Kamiya, A.,
 Kawai, J., Kim, C., Lin, J., Liu, S.X., Narusaka, M., Pham, P.K.,
 Sakano, H., Sakurai, T., Satou, M., Seki, M., Shinn, P., Yamada, K.,
 Shinzaki, K., Ecker, J., Theologis, A. and Davis, R.W.
 Direct Submission

TITLE
 JOURNAL
 Submitted (28-AUG-2001) DNA Sequencing and Technology Center,
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA

COMMENT
 e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAFL cDNAs (RAFL cDNA: "RIKEN
 Arabidopsis Full-Length cDNA") : Seki, M., Narusaka, M., Ishida, J.,
 Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
 Hayashizaki, Y. and Shinzaki, K.

The Salk, Stanford, PDEC (SSP) Consortium members carried out the
 sequencing and annotation of the RAFL cDNAs: Nguyen, M.,
 Southwick, A., Karlin-Neumann, G., Lam, B., Miranda, M., Palm, C.J.,
 Bowser, L., Jones, T., Banh, U., Chen, H., Cheuk, R., Chung, M.K.,
 Kim, C., Lin, J., Liu, S.X., Pham, P.K., Sakano, H., Shinn, P.,
 Yamada, K., Ecker, J., Theologis, A. and Davis, R.W.

Southwick, A., (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed
 equally to this work. Shinzaki, K. (RIKEN GSC) and Davis, R.W.
 (SSP/Stanford) contributed equally to this work as PIs.

FEATURES

source

1. 1938
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /chromosome="1"
 /clone="RAFL07-14-009"
 /note="This clone is in pluescript"
 1. 1938
 /gene="At1g49710; F14J22.8"
 160. 1701
 /gene="At1g49710; F14J22.8"
 /codon_start=1
 /product="putative fucosyltransferase"
 /protein_id="AAK6713.1"
 /db_xref="GI:15450884"
 /translation="MGVFSNLGRPRAGATHDERPATNGSPSSSSSSSIKRLSNL
 PLSALVLAIEIGFLRLDKVALVDLTDFEFTSPSLSSPPRSRRKRTGLFTDSC
 EEWLMREDSVYTSRDTKDFITISGEKIDQWCVCTFEDSGKTPDAFGLQKPG
 TSLIISMSDAQYPPENDLAQARRGYDVTMTSLSDVPVGFSAEVDIMSPVQPK
 TERAIAAEISNGARNFRLOALEALMKNIKIDSYGGRNDRGKVDKVEALKRYK
 SLAFENTNEDYVTEKEFOSILVAGSVVGPENIEEPFAASDFIHKITMEDVEEVA
 KRMVLAANDPAAYNOTLRMKYEGSDSPFALYNMAVHSCRCIFLATPREDEEES
 PNKRKPKCSKSGSDTYVHYVPREGRMESEVFLRGSYTDLEASVLAKEFKLK
 HEAVWKRKRPGLNKGDELRKIRHYIPGLITRQALYNFKREGSSLSHLQNPCKAF
 EVFV"
 *

gene

1. 1938
 /gene="At1g49710; F14J22.8"
 160. 1701
 /gene="At1g49710; F14J22.8"
 /codon_start=1
 /product="putative fucosyltransferase"
 /protein_id="AAK6713.1"
 /db_xref="GI:15450884"
 /translation="MGVFSNLGRPRAGATHDERPATNGSPSSSSSSSIKRLSNL
 PLSALVLAIEIGFLRLDKVALVDLTDFEFTSPSLSSPPRSRRKRTGLFTDSC
 EEWLMREDSVYTSRDTKDFITISGEKIDQWCVCTFEDSGKTPDAFGLQKPG
 TSLIISMSDAQYPPENDLAQARRGYDVTMTSLSDVPVGFSAEVDIMSPVQPK
 TERAIAAEISNGARNFRLOALEALMKNIKIDSYGGRNDRGKVDKVEALKRYK
 SLAFENTNEDYVTEKEFOSILVAGSVVGPENIEEPFAASDFIHKITMEDVEEVA
 KRMVLAANDPAAYNOTLRMKYEGSDSPFALYNMAVHSCRCIFLATPREDEEES
 PNKRKPKCSKSGSDTYVHYVPREGRMESEVFLRGSYTDLEASVLAKEFKLK
 HEAVWKRKRPGLNKGDELRKIRHYIPGLITRQALYNFKREGSSLSHLQNPCKAF
 EVFV"
 *

CDS

1. 1938
 /gene="At1g49710; F14J22.8"
 160. 1701
 /gene="At1g49710; F14J22.8"
 /codon_start=1
 /product="putative fucosyltransferase"
 /protein_id="AAK6713.1"
 /db_xref="GI:15450884"
 /translation="MGVFSNLGRPRAGATHDERPATNGSPSSSSSSSIKRLSNL
 PLSALVLAIEIGFLRLDKVALVDLTDFEFTSPSLSSPPRSRRKRTGLFTDSC
 EEWLMREDSVYTSRDTKDFITISGEKIDQWCVCTFEDSGKTPDAFGLQKPG
 TSLIISMSDAQYPPENDLAQARRGYDVTMTSLSDVPVGFSAEVDIMSPVQPK
 TERAIAAEISNGARNFRLOALEALMKNIKIDSYGGRNDRGKVDKVEALKRYK
 SLAFENTNEDYVTEKEFOSILVAGSVVGPENIEEPFAASDFIHKITMEDVEEVA
 KRMVLAANDPAAYNOTLRMKYEGSDSPFALYNMAVHSCRCIFLATPREDEEES
 PNKRKPKCSKSGSDTYVHYVPREGRMESEVFLRGSYTDLEASVLAKEFKLK
 HEAVWKRKRPGLNKGDELRKIRHYIPGLITRQALYNFKREGSSLSHLQNPCKAF
 EVFV"
 *

ORIGIN	BASE COUNT	559 a	380 c	430 g	569 t
Query Match	30.9%;				
Best Local Similarity	68.4%;				
Matches 989;	Conservative	0;	Mismatches 442;	Indels 15;	Gaps 3;
OY	301	AGAGAGAAATGAGACAAATCTAATGCGCTCTGTTGGTGGCCCTGGTGCGTCCGGGAGATC	360		
DB	268	AMCGCAAAATTTTCGAATTTGTTACCACTCTCCGTTCCTGTGTATATCCCTGGATC	327		
OY	361	GGGTTTCGGTAGGTGGATATGGCCAAAGAGCGGCCCATGGTGGACGCCCTCGCTGAC	420		
DB	328	GGGTTTCGGTAGGTGGATATGGCCAAAGAGCGGCCCATGGTGGACGCCCTCGCTGAC	378		
OY	421	TTCTTCTACCGCTCTCGAGCGGTCGTTGAGGTGACAGTATTTGGGCTTGGGTGGCT	480		
DB	379	TTCTTCTACCGCTCTCGAGCGGTCGTTGAGGTGACAGTATTTGGGCTTGGGTGGCT	438		
OY	481	TCGTATCGGAATTTGGAATCGTATGTTAGTTGGGAGATGGTTGAGAGGGAGATGCTGC	540		
DB	439	ATCGGATTTATTTACTGATAG--AGCTCGGAGAGTGGTTATGAGAAATTCAGTT	495		
OY	541	ACGATTTGAGAGGGGCTTTTCCAAAGAGCCCTATTTTGTTCGAGCTGATAGAGATGG	600		
DB	496	ACTTACTAGAGATTTTCTAAAGATCCAAATTTTATCTCGTGGGTGAAAGAGCATTT	555		
OY	601	AGTGTGTTGGGTTGATGTAATTTTGGGTTTGTGGGGATAGAAAGCCAGATCCGCA	660		
DB	556	CAATGGTGTCTGTGGATGTACATTTGGAGATGTTAGGGGAAACACAGATGCTCGC	615		
OY	661	TTTGGGTTTCTCAACCAAGTGAACAGTACATTTCTCGATTCATAGGAATCAGAGAA	720		
DB	616	TTTGGATTTAGGTGAGAAACCTGGAATCTTATATATATCTTCCATGGAAATCAGACAG	675		
OY	721	TACTATGCTGAGACAAATTTTCCATGGCAGACGAGGGATATACATCTGAATGACA	780		
DB	676	TATTTATCCAGAAATATGATTTCTCACAGGACGACGAGAGGTTATGATATGATGATC	735		
OY	781	ACCACTATATCTTGGAGTGTCTGTTGGATATTTTTCATGGGCTGATGATATGATG	840		
DB	736	ACTAGCTATCATCAGATGCTGCTGTGGATATTTTTCGTGGGCGAGATATATTTATG	795		
OY	841	GCACAGTGCACCGCAAAATGGAAGTGCCTGAGAGCTTTCATTTCCAAATGCTGT	900		
DB	796	TCTCGGATACACCAAAATGAGAGCTATTTGACGCTTTTATTTTCTATTTGTGCT	855		
OY	901	GCTCAAAATTTCCGGTGGCAGCTCTTGGAGCCCTTGAAAAATCAACATCAAAATTCAT	960		
DB	856	GCTCGGAATTTTCGTCTACAAACACTTGAAGCATTTGAAACATTAACATTAAGTATGAT	915		
OY	961	TCTTATGTTGTTGTACAGAACCGTGTGGAAGATGGAACAAGTGGAGCCCTGAGT	1020		
DB	916	TCTTATGTTGTTGTACAGAACCGGATGGGAAGTTGACAAGTTGAACCTCTTAAG	975		
OY	1021	CACCTCAAAATTTAGCTTACCGTTGAAAATTTGAGATGAGAGAAGATTTATTAACCTGAAAA	1080		
DB	976	CGATCAAAATTTAGTTTGGCTTTTATAGATTAACAGAGAGATTTATCTCCCGAGAA	1035		
OY	1081	TTCTTCAATCCCTTGTGCTGGAAGTGCCTGTTGGTGTGGTGGTGCCTCAATATTGAG	1140		
DB	1036	TTCTTCAATCCCTTGTGCTGGAAGTGCCTGTTGGTGTGGTGGTGCCTCAATATTGAG	1095		
OY	1141	GACTTGTCTCTTCTCGTTCAATTTTACATATTAAGAGATAGAGATGTTGAGTCT	1200		
DB	1096	GAAATTTGGCGCTCTCGGACATTCCTCTCAATTTAAGATATGGAAGATGTAGAGCCA	1155		
OY	1201	GTTGCAAGACCATGATATCTACAGAAAAATCCGGAACATATATATCATTCATTTGAGG	1260		
DB	1156	GTTGCAAGACCATGATATCTACAGAAAAATCCGGAACATTCATATATCATTCATTTGAGG	1215		
OY	1261	TGGAGTATGAGGGTCAATCTGATCTTCAAGGCCCTTGTGGATATGGCACCTGTGAT	1320		

Db	1216	TGGAATAACGAGGGGTCCTTCAGATTCTTTCAAGGACACTGTTGATATGCGTGTACAC	1275
QY	1321	TCATCGTGCCTCTTTTGCATTCACCTTGCCCAACAGTAGAGAGAGAGAAAATAT	1380
Db	1276	TCTTCTTGCCCTCTCTGCAATTTTCTTGCCCAAGAGGTCGCCAAGAGAGAAAGC	1335
QY	1381	CCAAGCCTTAGAGACGTCCTTGCAAGT---GCACTAGAGGGCCAGAAACGTTATCAT	1437
Db	1336	CCTAATTTCAAGAAACACGCCCTGCCAATGTAGCAGGGAGATCAGACACAGTTATCAT	1395
QY	1438	ATCATGTGACAGAAAGGGGAAAGTTTGATGAGTGCATTTACCTGAGGCTTAGCAAT	1497
Db	1396	GTTTTTGTAGAGAAAGAGCGCGGTTTGAAATGCAATCAGTCTTTTGAAGGATAAAGT	1455
QY	1498	TTAACTGCAATGCTGTGTAAGCGCTGCTTTGTTTGAAGTTCACATCCCTCAATCTGTG	1557
Db	1456	GTGACTCAGGAAGCTCTAGAAATCTGCACCTTCTCGCAAGTCTCAAGTCTTTAAACATGAG	1515
QY	1558	CCTGATGAGAGACTGAAAGGCTCGAACTTTPAAGAGGGGGAGTGTCTTAAACTGAC	1617
Db	1516	GCACTGTGGAAGAGAGAAAGCGCTGGAACTTAAAGAGACAAAGAGCTTAAATACAT	1575
QY	1618	AAATATATCCCAATTTGCTTGACACAGAGACAGCTCTTTATACCTTCAGCTTCAAGGT	1677
Db	1576	CGGATTTACCCGCTTGGCTTAACGGAAGACAGAGGTTTGTCAACTCAAAATGAGAGGA	1635
QY	1678	GATGCTGATTTTCAGAGCTCACTTGGAGACAATCTTGTGCCAAGTTGGAAGTCATTTT	1737
Db	1636	AATTCGACTCTAAGTAGTACATTCACAAACCAACCCCTTGCTCAATTTGAGGTGTCTTC	1695
QY	1738	GTGTAG 1743	
Db	1696	GTCTAG 1701	
RESULT 7	ATHA04861		
LOCUS	ATHA04861		
DEFINITION	Arabidopsis thaliana mRNA for truncated alpha1,3-fucosyltransferase (FucTB gene).		
ACCESSION	AJ404861		
VERSION	AJ404861.2		
KEYWORDS	GI:21912393		
SOURCE	alpha 1,3-fucosyltransferase; FucTB gene.		
ORGANISM	Arabidopsis thaliana.		
REFERENCE	Arabidopsis thaliana.		
REFERENCE	1		
REFERENCE	Wilson, I.B., Rendic, D., Frellinger, A., Dumlac, J., Altmann, F., Mucha, J., Muller, S. and Hauser, M.T.		
REFERENCE	Cloning and expression of cDNAs encoding alpha1,3-fucosyltransferase homologs from Arabidopsis thaliana		
REFERENCE	Biochim. Biophys. Acta 1527 (1-2), 88-96 (2001)		
REFERENCE	2		
REFERENCE	Wilson, I.B.H.		
REFERENCE	Direct Submission		
REFERENCE	Submitted (11-JUL-2000) Wilson I.B.H., Institut fuer Chemie, Universitaet fuer Bodenkultur, Muthgasse 18, Wien 1190, Austria		
REFERENCE	3 (bases 1 to 1834)		
REFERENCE	revised by [4]		
REFERENCE	Wilson, I.B.H.		
REFERENCE	Direct Submission		
REFERENCE	Submitted (10-SEP-2001) Wilson I.B.H., Institut fuer Chemie, Universitaet fuer Bodenkultur, Muthgasse 18, Wien 1190, Austria		
REFERENCE	On Jul 19, 2002 this sequence version replaced gi:1592484.		
REFERENCE	Location/Qualifiers		
REFERENCE	1..1834		
REFERENCE	/organism="Arabidopsis thaliana"		
REFERENCE	/strain="Columbia"		
REFERENCE	/db_xref="taxon:3702"		
REFERENCE	/clone_id="Lambda Zap II cDNA"		


```

gene
  /country="Portugal"
  2..658
  /gene="fuct3"
  <2..>658
  /gene="fuct3"
  /codon_start=1
  /evidence=experimental
  /product="alpha1,3-fucosyltransferase"
  /protein_id="CAC95159.1"
  /db_xref="GI:16304828"
  /translation="ENNIMARRRGVHIAMTSSIDNPVGYFSMAEYDIMPSTPT
  EKAAAFISNCGARNFLOALELKTNTSIDYSGCHNRDRVRLKLTIRKFS
  LAENSNEDYVTEKFEQSLVAGTIPVVGPPNIDPASPSPFLYIKLEDESVAK
  SMRYLAENPEAYNOSLRMKYEGSDSFKALVDMAVXHSRCRLCIHLAKSREKEKSP
  D"
BASE COUNT      188 a      135 c      149 g      185 t      1 others
ORIGIN
Query Match      24.3%  Score 534;  DB 8;  Length 658;
Best Local Similarity 88.4%;  Pred. No. 4,2e-125;
Matches 579;  Conservative 0;  Mismatches 76;  Indels 0;  Gaps 0;

OY  729  TGAACAATATTTGCGATGCGAAGAGGAGGATATACATCGAATGACACACAGTCT 788
      |||||||
DB   1  TGAACAATATTTGCGATGCGAAGAGGAGGATATCAATGCAATGACACACAGTCT 60
      |||||||

OY  789  ATCTTCGATGTCCTGCTGTGGATATTTTTCATGGGCTGATGATGATGATGACACAGT 848
      |||||
DB   61  ATCATCGACGTCCTGCTGTGGATATTTTTCATGGGCTGATGATGATGATGACACAGT 120
      |||||

OY  849  GCAGCCGAAACTGAGCTGCTTTCAGCTGCTTTCATTTCCATTTGCTGCTCGAAA 908
      |||||
DB   121  AAGCCGAAACTGAAAGCTTTTTCAGCTGCTTTCATTTCCATTTGCTGCTCGAAA 180
      |||||

OY  909  TTTCCGGTTGAGCTGCTTTCAGCTGCTTTCAGCTGCTTTCATTTCCATTTGCTGCTG 968
      |||||
DB   181  TTTCCGGTTGAGCTGCTTTCAGCTGCTTTCAGCTGCTTTCATTTCCATTTGCTGCTG 240
      |||||

OY  969  TGGTTGTCACGAAACCGTGTAGAGAGAGTGAACAAAGTGAAGCCCTGAGCACTACAA 1028
      |||||
DB   241  TAGTTGTCATGGAATGCTGTAGAGAGAGTGAACAAAGTGAAGCCCTGAGCACTACAA 300
      |||||

OY  1029  ATTTAGCTTACGCTTTGAAATTCGATGAGAAATTTATTAACGTAAATTTCTTCCA 1088
      |||||
DB   301  ATTTAGCTTACGCTTTGAAATTTCTAAGAGAGATTTATTAACGTAAATTTCTTCCA 360
      |||||

OY  1089  ATCCCTGTTGCTGGAACGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1148
      |||||
DB   361  GTCCCTGTTGCTGGAACGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
      |||||

OY  1149  TCCCTTCTCTGCTCAATTTTATATTAAGAGATGAGAGATGTTGAGTCTGTTGCAA 1208
      |||||
DB   421  TCCCTTCTCTGCTCAATTTTATATCAAAAGAACTGAGAGATGTTGAGTCTGTTGCAA 480
      |||||

OY  1209  GACCATGAGATTTCTAGCAGAAATCCGAAAGCATATATATCATTTGAGGTGGAAGTA 1268
      |||||
DB   481  GTCATGAGATTTCTAGCAGAAATCCGAAAGCATATATATCATTTGAGGTGGAAGTA 540
      |||||

OY  1269  TGAGGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1328
      |||||
DB   541  TGAAGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
      |||||

OY  1329  CCGTCTTTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1383
      |||||
DB   601  CCGGCTTTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 655
      |||||

RESULT 9
AF277228      787 bp      mRNA      linear      PLN 29-JUN-2002
LOCUS        Arabidopsis thaliana clone 3M alpha 1,3-fucosyltransferase mRNA,
DEFINITION   partial cds.
ACCESSION   AF277228

```

```

VERSION      AF277228.1  GI:21628751
KEYWORDS
SOURCE       thale cress.
ORGANISM     Arabidopsis thaliana
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE    1 (bases 1 to 787)
AUTHORS      Kiefer-Meyer, M.-C., Faye, L. and Gomord, V.
TITLE        Direct Submission
JOURNAL      Submitted (08-JUN-2000) LTI, CNRS UMR 6037, Faculte des Sciences,
              Universite de Rouen, Mont Saint Aignan 76821, France
FEATURES
  source
    /organism="Arabidopsis thaliana"
    /db_xref="taxon:3702"
    /clone="3M"
    /note="ecotype: WS"
    <1..>787
    /codon_start=2
    /product="alpha 1,3-fucosyltransferase"
    /protein_id="FAM68912.1"
    /db_xref="GI:21628752"
    /translation="STSSSSPSSSEFKKRVSTPLPCVALVYIEIGFLRLDNLAVLD
    TLTHFFKSSSDLKVGSGIEKCEWELRVDSVYSRDPDLPFIISGNDFKSCVD
    CVMGFTSDKRDPAFLSHQPTLSIIRSESAYVOENMLAORRRGYDIVTTSLS
    SDVPGFSMAEYDIMPVOPKTEELAAAFISNCGARNFLOALELAKMTNWKIDSY
    GGCHRNDSGVEVEALKHKFSLAEPENTNEEDYVTEKFEQSLV"
BASE COUNT   202 a      160 c      178 g      247 t
ORIGIN
Query Match      15.9%  Score 349;  DB 8;  Length 787;
Best Local Similarity 73.6%;  Pred. No. 6.4e-78;
Matches 445;  Conservative 0;  Mismatches 160;  Indels 0;  Gaps 0;

OY  493  TCTGATCTGTAAGTTGAGGAATGTTGGAGAGGAGATGCTGTCAGCTATTGAGG 552
      |||||
DB   182  TCAGGAATGAGAAATGCGAGAGTGTGAGAGAGGAGATTCAGTACTTATTTCTAGA 241
      |||||

OY  553  GGCCTTTCGAAAGAGCCATTTTGTCTGAGCTGTCAGAGTGGAGTGGAGTGGTTCG 612
      |||||
DB   242  GATTTCCTAAAGATCCGATTTTATCTCTGTTGTAACAAGCACTTCAAAATGCTGCT 301
      |||||

OY  613  GTTGATGTAAATTTGGTTAGTGGGATGAGAAAGCCAGATGCCGATTTGGTACTCT 672
      |||||
DB   302  GTTGATGTGTAATGGATTCACATTCAGATTAAGAAACCTGATGGGCTTTGCATTAAGT 361
      |||||

OY  673  CAACCACTGGAACAGCTAGCATTTCTGATCAATGCAATGAGACAGAACTACTATGCTGAG 732
      |||||
DB   362  CATCAACCTGGAACAGCTAGCATTTCTGATCAATGCAATGAGACAGAACTACTATGCAAGAG 421
      |||||

OY  733  AACATATTGCTAGGAGAGAGAGGAGATTAATCATGTAATGCAACAGCTACTACT 792
      |||||
DB   422  AATATCTTCTGCTCAAGCAGCAGGAAAGGTTATATATTGATGACAACTAGCTGTCA 481
      |||||

OY  793  TCGATGTTCTGTTGGATATTTTTCATGGCTGAGATGATATGATGACACAGTCCAG 852
      |||||
DB   482  TCAATGTTCTGTTGGATATTTTTCATGGGCGGAATATGATATTAATGCTCCAGTCAA 541
      |||||

OY  853  CCGAAACTGAGAGCTGCTTTCGAGCTGCTTTCATTTCCATTTGCTGCTGCAAAATTC 912
      |||||
DB   542  CCAAAACAGAGAGAGCTTTCGAGCTGCTTTCATTTCCATTTGCTGCTGCAAAATTC 601
      |||||

OY  913  CGGTTCAGAGCTTTCGAGAGCTTTCGAAATCAAAATCAAAATTTGATTTTATGTTGGT 972
      |||||
DB   602  CGCCTGCAAGCTCTTGAAGCTTTAATGAGACGATGTTAAGATTTGATTTATGTTGGT 661
      |||||

OY  973  TGTCAAGAGAAACGATGAGAGAGTGAACAAATGAGAGCCCTGGAAGCACTACAAATTT 1032
      |||||
DB   662  TGTCAAGAGAAACGATGAGAGAGTGAACAAATGAGAGCCCTGGAAGCACTACAAATTT 721
      |||||

OY  1033  AGCTTAGCGTTGAAATTTCAATGAGAGAGATTTATGTAACGAAATTTCTTCAATCC 1092
      |||||

```

DB	722	AGTCAGCTTTTGAGAACACCAACGAGGAGGATTTGTGCACAGAGAAGTTCTTCCAACT	781
QY	1093	CTTGT 1097	
Db	782	CTAGT 786	
RESULT 10			
LOCUS	AF277229	588 bp	mRNA linear
DEFINITION	Arabidopsis thaliana clone 1F alpha 1,3-fucosyltransferase mRNA,		
ACCESSION	AF277229		
VERSION	AF277229.1	GI:21628753	
KEYWORDS			
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta; Spermatophyta: Magnoliophyta: eudicotyledons; core eudicots; Rosidae; eurosids II: Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 588)		
AUTHORS	Kiefer-Meyer,M.-C., Faye,L. and Gomord,V.		
TITLE	Direct Submission		
JOURNAL	Submitted (08-JUN-2000) LRI, CNRS UMR 6037, Faculte des Sciences, Universite de Rouen, Mont Saint Aignan 76821, France		
FEATURES			
source	Location/Qualifiers		
	1..588		
	/organism="Arabidopsis thaliana"		
	/db_xref="taxon:3702"		
	/clone="11p"		
	/note="ecotype: WS"		
	<1..>588		
	/codon_start=1		
	/product="alpha 1,3-fucosyltransferase"		
	/protein_id="AA068913.1"		
	/db_xref="GI:21628754"		
	/translation="QEWLEKDYSDYTSRDPFTKDPFIISGSKDKFSCSVDVCWMEFTSD KRPDAARGLSHQPTGLTSLIRMSESAQYQYENNNLAQARKGDIWMTLSLSDVGVGF SWAPYDIMAPOPTERKALAAAFISNCAARNFRLQALEFKMTWKIDSYGCHRNDD GSVEKVEALKHFKSLAEFNENEDDYTEKPFQSLV"		
BASE COUNT	167 a 108 c 137 g 176 t		
ORIGIN			
Query Match	15.7%;	Score 346;	DB 8; Length 588;
Best Local Similarity	74.4%;	Pred. No. 3.7e-77;	
Matches 436;	Conservative 0;	Mismatches 150;	Indels 0; Gaps 0;
QY	512	AGGAATGTTGGAGAGGAGATGCTGTACACTATTTCGAGGGGCTTTCCAAAGGCTTA	571
Db	2	AGGAATGTTGGAGAGGAGATGCTGTACACTATTTCGAGGGGCTTTCCAAAGGCTTA	61
QY	572	TTTTTTGTTTGAGACTGATCAGAGTGGAGATCGTGTGGTGGTGGATTAATTTGGAT	631
Db	62	TTTTTATCTCTGTAAGTAACAAGACTTCAATCTGCTCTGTATTTGTGTAATGGCAT	121
QY	632	TTTAGTGGGATGAAGAAGCCAGATGCGGATTTGGGTTACCTCAACCAAGTGAACAGCTA	691
Db	122	TCACCTTCAGATTAAGAAACCTGATGCGGCTTTGGATTAAAGTCATCAACCTGGAACATCA	181
QY	692	GCATTCCTCGCATCATGCAATTCAGCAAGATTAATGCTGAGACAAATATGGCATGGCAA	751
Db	182	GTAATATCCGTTCCATGGAATCAGCAGTATTAACCAAGAGATTAATCTTGCTCAAGCAC	241
QY	752	GACGAGGGGATTAATACATCGTAATGACAAACAGTCATCTTCGGATTTCTCTGTGGAT	811
Db	242	GACGGAAGGATTAATATTTGTGATGACAACTAGTCTGTCAATCAATGTCTCTGTGGAT	301
QY	812	ATTTTTCATGGCGGAGATGATGATGATGAGCAACAGTCCAGCCGGAAGATGAAGTGGTC	871
Db	302	ATTTTTCATGGCGGAGATGATGATGATGAGCGTCCAGTCCAAACCAAGAGAAAGCTC	361
QY	872	TTGACGCTGCTTTCATTTCCAAATGTGGTGCTCGAAATTTCCGGTTGCAAGCTCTTGAGG	931

LOCUS	AX105537	2069 bp	DNA	Linear	PAT 30-APR-2001
DEFINITION	Sequence 3 from Patent WO0125454.				
ACCESSION	AX105537				
VERSION	AX105537.1	GI:13921581			
KEYWORDS					
SOURCE	synthetic construct.				
ORGANISM	artificial sequences.				
REFERENCE	1 (bases 1 to 2069)				
AUTHORS	V Zina, L.P. and D'Aoust, M.A.				
TITLE	Method for regulating transcription of foreign genes in the presence of nitrogen				
JOURNAL	Patent: WO 0125454-A 3 12-APR-2001;				
FEATURES	Medicago Inc. (CA)				
source	Location/Qualifiers				
	1..2069				
	/organism="synthetic construct"				
	/db_xref="taxon:32630"				
	/note="Sequences to be used as promoter for regulating expression of foreign genes"				
BASE COUNT	640 a 389 c 335 g 705 t				
ORIGIN					
Query Match	9.5%	Score 209.4;	DB 6;	Length 2069;	
Best Local Similarity	66.5%;	Pred. No. 2.8e-42;			
Matches 341;	Conservative 0;	Mismatches 151;	Indels 21;	Gaps 2;	
Oy 1486	AGGCTAGCAATTTAACTGTAAGCTGGAAGGCTGCTGTTGTTGAAGTTCACATCC	1545			
Db 170	AGATCAGCAATTTAACTGTAAGCTGGAAGGCTGCTGTTGTTGAAGTTCACATCC	229			
Oy 1546	CTGAATCTTGCTCCGTATGGAAGAAGGCTGGAAGTATTAAGAAGGGGAGTCT	1605			
Db 230	CTGAATCTTGCTCCGTATGGAAGAAGGCTGGAAGTATTAAGAAGGGGAGTAA	289			
Oy 1606	TTAAACGCTCAAAATTTATCCCAATTTGGCTTGACACAGACAGCTTATACCTTC	1665			
Db 290	TTGAAGGTTTAAATAATATACCTCGGGGCTGGAACAGAGGCAAGCTTATACCTTC	349			
Oy 1666	AGCTTCAAGGATGATGATTTCAAGATCCTTGGAAGCAATCCTTGTCGAAGTTT	1725			
Db 350	CAGTTCACAGGGGATGTTGATTTCAAGATCCTTGGAAGCAATCCTTGTCGAAGTTT	409			
Oy 1726	GAAGTCAATTTTGTGTAGCATGCCGCTAA-----ATGGTACCTCTGCTACCTGAATT	1778			
Db 410	GAAGTCAATTTTGTGTAGCATGCCGCTAA-----ATGGTACCTCTGCTACCTGAATT	469			
Oy 1779	AGCTTCACTTACGTCAGCACTAGTAGAGTTTGAAGATGAGTATGCGATGCAATATGCG	1838			
Db 470	AGCTTCACTTACGTCAGCACTAGTAGAGTTTGAAGATGAGTATGCGATGCGC	529			
Oy 1839	ATGGCTT-----TATTATGCGAGTTTCTTGCGCAACATTCATGATGTTT	1884			
Db 530	ATGGCTTGTATAGCTCCTACCTACTTTGGCCACACTCATCGGGGATTTTCATTACAGAAAT	589			
Oy 1885	GTAATAGCATCACACTTAAATTTTAAACTGTGTTCTGTAGAGAGTGCAAATCCATATTTA	1944			

KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 2232)
AUTHORS v Zina, L.P. and D'Aoust, M.A.
TITLE Method for regulating transcription of foreign genes in the presence of nitrogen
JOURNAL Patent: WO 0125454-A 12 12-APR-2001;
Medicago Inc. (CA)
FEATURES
Source location/Qualifiers
1. 2232
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Sequences to be used as promoter for regulating expression of foreign genes"
BASE COUNT 686 a 436 c 362 g 748 t
ORIGIN
Query Match
Best Local Similarity 66.5%; Pred. No. 2.8e-42;
Matches 341; Conservative 0; Mismatches 151; Indels 21; Gaps 2;
QY 1486 AGCTTACGAATTTAACTGCAATGCTGTGAAGGCTGCTGTTGTTGAAGTTCACATCC 1545
DB 170 AGATCTAGCAATTTAACTGCAATGCTGTGAAGGCTGCTGTTGTTGAAGTTCACATCC 229
QY 1546 CTGATCTTGCTGTATGGAAGCACTGAAAGGCTGCAAGTTAAGAGGGGGAGTGCCT 1605
DB 230 CTGATCTATGCTGTATGGAAGCACTGAAAGGCTGCAAGTTAAGAGGGGGAGTGCCT 289
QY 1606 TTAAGCTACAAATAATATACCAATTTGCTGTGACAGAGCAAGCTTTATACCTTC 1665
DB 290 TTGAAGTTTACAAATAATATACCAATTTGCTGTGACAGAGCAAGCTTTATACCTTC 349
QY 1666 AGCTTCAAGGTATGCTGTATGGAAGCACTGTAAGCAATCTTGTGCCAAGTTT 1725
DB 350 CAGTTCAACGGGATGTTGATTCAGAAATCATTGGAAGCAATCTTGTGCCAAGTTT 409
QY 1726 GAAGTCATTTTGTGTGATGCGCTAA-----ATGATACCTCTGCTACCTGATTT 1778
DB 410 GAAGTCATTTTGTGTGATGCGCTAA-----ATGATACCTCTGCTACCTGATTT 1778
QY 1779 AGCTTCACTAGTACGACACTAGCTAGATTTTGAAGATGATGCGACATGATATGCG 1838
DB 470 AGCTTCACTAGTACGACACTAGATTTTGAAGATGATGCGACATGATATGCG 529
QY 1839 ATGGCTT-----TATTTATGCTAGTTCTTGCCCACTGATGATGTTT 1884
DB 530 ATGGCTTGTATGCTAGTACCTTGTGCGCACTCATCGGGGATTTTACATTCAGAAAT 589
QY 1885 GTATAGACATCACACTTATATTTTAACCTGTTCTGTAAGATGCAATCCATATTTA 1944
DB 590 ATACATGACTTCAACCACTTAAACCCCTTTTGTGAAGATTAAGTATGTTATATTTA 649
QY 1945 ATGCTAGTTTATGCTCTTATCTGATCATCT 1977
DB 650 ATGTTGGTTGTAGTGTTTTACTGATTAAT 682

RESULT 15
AX105536 2808 bp DNA linear PAT 30-APR-2001
LOCUS AX105536
DEFINITION Sequence 2 from Patent WO0125454.
ACCESSION AX105536
VERSION AX105536.1 GI:13921580
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 2808)
AUTHORS v Zina, L.P. and D'Aoust, M.A.
TITLE Method for regulating transcription of foreign genes in the

JOURNAL Presence of nitrogen
Patent: WO 0125454-A 12-APR-2001;
Medicago Inc. (CA)
FEATURES
Source location/Qualifiers
1. 2808
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Sequences to be used as promoter for regulating expression of foreign genes"
BASE COUNT 879 a 502 c 481 g 946 t
ORIGIN
Query Match
Best Local Similarity 66.5%; Pred. No. 2.8e-42;
Matches 341; Conservative 0; Mismatches 151; Indels 21; Gaps 2;
QY 1486 AGCTTACGAATTTAACTGCAATGCTGTGAAGGCTGCTGTTGTTGAAGTTCACATCC 1545
DB 909 AGATCTAGCAATTTAACTGCAATGCTGTGAAGGCTGCTGTTGTTGAAGTTCACATCC 968
QY 1546 CTGATCTTGCTGTATGGAAGCACTGAAAGGCTGCAAGTTAAGAGGGGGAGTGCCT 1605
DB 969 CTGATCTATGCTGTATGGAAGCACTGAAAGGCTGCAAGTTAAGAGGGGGAGTGCCT 1028
QY 1606 TTAAGCTACAAATAATATACCAATTTGCTGTGACAGAGCAAGCTTTATACCTTC 1665
DB 1029 TTGAAGTTTACAAATAATATACCAATTTGCTGTGACAGAGCAAGCTTTATACCTTC 1088
QY 1666 AGCTTCAAGGTATGCTGTATGGAAGCACTGTAAGCAATCTTGTGCCAAGTTT 1725
DB 1089 CAGTTCAACGGGATGTTGATTCAGAAATCATTGGAAGCAATCTTGTGCCAAGTTT 1148
QY 1726 GAAGTCATTTTGTGTGATGCGCTAA-----ATGATACCTCTGCTACCTGATTT 1778
DB 1149 GAAGTCATTTTGTGTGATGCGCTAA-----ATGATACCTCTGCTACCTGATTT 1208
QY 1779 AGCTTCACTAGTACGACACTAGCTAGATTTTGAAGATGATGCGACATGATATGCG 1838
DB 1209 AGCTTCACTAGTACGACACTAGATTTTGAAGATGATGCGACATGATATGCG 1268
QY 1839 ATGGCTT-----TATTTATGCTAGTTCTTGCCCACTGATGATGTTT 1884
DB 1269 ATGGCTTGTATGCTAGTACCTTGTGCGCACTCATCGGGGATTTTACATTCAGAAAT 1328
QY 1885 GTATAGACATCACACTTATATTTTAACCTGTTCTGTAAGATGCAATCCATATTTA 1944
DB 1329 ATACATGACTTCAACCACTTAAACCCCTTTTGTGAAGATTAAGTATGTTATATTTA 1388
QY 1945 ATGCTAGTTTATGCTCTTATCTGATCATCT 1977
DB 1389 ATGTTGGTTGTAGTGTTTTACTGATTAAT 1421

Search completed: April 28, 2003, 02:56:19
Job time : 5634 secs

